Master’s Thesis

Setup of human naïve pluripotency reversion protocol to study the function of early embryonic genes

Joonas Sokka

Master's Programme in Genetics and Molecular Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki

October 2019

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Faculty	Degree Programme
Faculty of Biological and Environmental Sciences	Master's Programme in Genetics and Molecular Biosciences

Author

Juho Joonas Sokka

Title

Setup of human naïve pluripotency reversion protocol to study the function of early embryonic genes

Subject/Study track

Cell and Developmental Biology

Level	Month and year	Number of pages
Master's Thesis	October 2019	64

Abstract

Pluripotent stem cells (PSC) can exists in both primed and naïve states. The conventionally derived human PSCs represent the later primed state of pluripotency during embryo development, while the naïve state resembles the inner cell mass (ICM) of pre-implantation blastocyst. Primed human PSCs can be reverted chemically by transient histone deacetylase (HDAC) inhibition back to the naïve state *in vitro*. The reverted PSCs can then be characterized based on their morphology and expression of selected naïve markers using immunocytochemistry and RT-qPCR assays.

Leucine twenty homeobox (LEUTX) is one of the genes expressed during the early stages of embryo development and is capable of activating the transcription of multiple genes, including pluripotency-associated genes, which are upregulated during the human embryonic genome activation (EGA). LEUTX expression could potentially improve the naïve reversion efficiency or the maintenance of naïve PSCs by driving the transcriptome of primed PSCs back towards the earlier cell stages of embryo development, potentially even to cell stages that precede the naïve state.

The aim of this thesis was to setup the naïve reversion protocol and study the effects of LEUTX on the reversion by using the generated and tested H9 activator cell line for targeted activation of endogenous LEUTX expression. First, a conditionally stabilized CRISPRa activator cell line was generated for targeted activation of endogenous gene expression in H9 cells. Then sequence-specific guide RNAs (gRNA) targeting LEUTX for activation were introduced to the activator cell line. Using the generated activator cell line during the naïve reversions allows the targeted activation of specific genes, here LEUTX, and thus enables studying the effects of these genes on PSCs during the naïve reversion protocol.

The induced activator cells expressing LEUTX managed to form four times as many naïve resembling colonies during the reversion compared to the controls, but most of these were lost after changing the medium conditions towards the end of the protocol. After the reversion was complete, the reverted PSCs were characterized as naïve PSCs based on their domed morphology and the high expression of naïve markers NANOG, KLF17, TFCP2L1 and DNMT3L when compared to the primed PSCs.

The naïve reversion protocol was set up and optimized successfully and can now be used as a reliable way of obtaining human naïve PSCs for further experiments studying and modelling the earlier developmental stages during embryo development. Furthermore, the generated H9 activator cell line worked as intended and can be utilized for studying the effects of other targeted genes during the reversion or in the reverted naïve PSCs.

Keywords

Primed pluripotent stem cells, Naïve pluripotent stem cells, Naïve reversion, LEUTX, ESCs, iPSCs, CRISPRa, CRISPR/Cas

Supervisor or supervisors

Ras Trokovic, Eeva-Mari Jouhilahti, Jere Weltner and Timo Otonkoski

Where deposited
HELDA - Digital Repository of the University of Helsinki

Additional information

Tiedekunta	Koulutusohjelma
Bio- ja Ympäristötieteellinen Tiedekunta	Genetiikan ja molekulaaristen biotieteiden maisterinohjelma

Tekijä

Juho Joonas Sokka

Työn nimi

Ihmisen naiivien pluripotenttien kantasolujen reversio protokollan valmistelu varhaisen alkionkehityksen geenien tutkimista varten

Oppiaine/Opintosuunta

Solu- ja Kehitysbiologia

Työn laji	Aika	Sivumäärä
Maisterin tutkielma	Lokakuu 2019	64

Tiivistelmä

Pluripotentit kantasolut (PS-solut) voivat esiintyä sekä erilaistumiseen alustetussa (primed) sekä naiivissa olomuodossa. Perinteisesti eristetyt ihmisen PS-solut edustavat myöhempää primed-tilaa alkionkehityksessä, kun taas naiivit PS-solut muistuttavat preimplantaatio vaiheen alkiorakkulan sisäsolumassaa (ICM). *In-vitro* tutkimuksissa ihmisen primed-PS-solut voidaan reversoida kemiallisesti tilapäisellä histoni deasetylaasin (HDAC) inhiboinnilla takaisin kohti aiempaa naiivia tilaa. Reversoidut PS-solut voidaan karakterisoida niiden morfologian sekä valittujen naiivi markkereiden ekspression perusteella käyttämällä immunosolukemiallisia sekä RT-qPCR testejä.

Leusiini 20 homeoboksi (LEUTX) on yksi ensimmäisistä ekspressoiduista geeneistä varhaisessa alkionkehityksessä, ja se kykenee aktivoimaan transkription useissa geeneissä, jotka ovat ylläpidettyjä ihmisen alkion genomin aktivaation (EGA) aikana. Näihin ylläpidettyihin geeneihin mukaan lukeutuu solujen pluripotenttisuuteen yhdistettyjä geenejä. LEUTX ekspressio voisi mahdollisesti parantaa naiivi reversion tehokkuutta tai edesauttaa naiivien PS-solujen ylläpitoa ajamalla primed-PS-solujen transkriptomia kohti alkionkehityksen aiempia soluvaiheita, mahdollisesti jopa naiivia tilaa edeltäviin soluvaiheisiin.

Tämän tutkielman tavoitteena oli valmistella naiivi reversio protokolla käyttöä varten ja samalla tutkia LEUTX:n vaikutusta naiivi reversioon käyttämällä valmistettua ja testattua H9 aktivaattorisolulinjaa kohdennettuun endogeenisen LEUTX ekspression aktivoimiseen. Ensin valmistettiin ehdollisesti stabiloitu CRISPRa aktivaattorisolulinja H9 soluilla kohdennettua endogeenisen geenieksprression aktivointia varten. Tämän jälkeen aktivaattorisoluihin tuotiin LEUTX-kohdentavia sekvenssispesifisiä opas-RNA:ita (gRNA). Käyttämällä valmistettua aktivaattorisolulinjaa naiivi reversioiden aikana voidaan kohdennetusti aktivoida spesifisiä geenejä, tässä tapauksessa LEUTX, mikä mahdollistaa niiden vaikutuksien tutkimisen PS-soluissa naiivi reversion aikana.

Indusoidut aktivaattorisolut, joissa LEUTX oli ekspressoituna, onnistuivat reversion aikana muodostamaan neljä kertaa enemmän naiivia tilaa muistuttavia kolonioita kontrolleihin verrattuna, mutta suurin osa näistä kolonioista ei selvinnyt kasvatusmedian vaihduttua reversio protokollan lopulla. Kun reversio saatiin valmiiksi, reversoidut PS-solut karakterisoitiin naiiveiksi PS-soluiksi niiden kupolimaisen morfologian sekä primed-PS-soluihin verrattuna korkeiden naiivimarkkereiden (NANOG, KLF17, TFCP2L1 ja DNMT3L) ekspressiotasojen perusteella.

Naiivi reversion protokolla saatiin valmisteltua ja optimoitua onnistuneesti, ja sen avulla voidaan jatkossa tuottaa luotettavalla tavalla ihmisen naiiveja PS-soluja tulevia tutkimuksia ja alkionkehityksen aikaisempien vaiheiden mallinnusta varten. Lisäksi valmistettu H9 aktivaattorisolulinja toimi odotetusti, ja sitä voidaan hyödyntää jatkossa valittujen geenien ekspressioiden vaikutuksien tutkimiseen reversion aikana sekä jo reversoiduissa naiivi PS-soluissa.

Avainsanat

Primed pluripotentit kantasolut, Naiivit pluripotentit kantasolut, Naiivi reversio, LEUTX, ES-solut, iPS-solut, CRISPRa, CRISPR/Cas

Ohjaaja tai ohjaajat

Ras Trokovic, Eeva-Mari Jouhilahti, Jere Weltner ja Timo Otonkoski

Säilytyspaikka
HELDA - Helsingin yliopiston digitaalinen arkisto

Muita tietoja

Table of Contents

Abbreviations

AP	Alkaline phosphatase
BMP	Bone morphogen protein
Cas	CRISPR associated protein
cDNA	Complementary DNA
CRISPR	Clustered regularly interspaced short palindromic repeat
CRISPRa	CRISPR activation
dCas9	Dead Cas9
DD	Destabilization domain
DHFR	Dihydrofolate reductase
DMSO	Dimethyl sulphoxide
DOX	Doxycycline
DNMT3L	DNA methyltransferase 3 like
E6	Essential 6 medium
E8	Essential 8 medium
EB	Embryoid body
EDTA	Ethylenediaminetetraacetic acid
EpiSC	Epiblast stem cell
ESC	Embryonic stem cell
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
FOXA2	Forkhead box protein A2
GFP	Green fluorescent protein

gRNA	Guide RNA
GSK3	Glycogen synthase kinase-3
H3K27me3	Histone protein H3 lysine 27 tri-methylation
HDAC	Histone deacetylase
HDACi	Histone deacetylase inhibitor
HFF	Human foreskin fibroblasts
ICM	Inner cell mass
iMEF	Inactivated mouse embryonic fibroblast
KLF17	Krüppel-like factor 17
LEUTX	Leucine twenty homeobox
LIF	Leukemia inhibitory factor
MAPK/ERK	Mitogen-activated protein kinase
MEK	Mitogen-activated protein kinase kinase
NEB5	New England Biolabs 5α transformation competent *E. coli* cells
NaB	Sodium butyrate
iPSC	Induced pluripotent stem cell
OCT4	Octamer-binding transcription factor 4
PB	Piggybac
PFA	Paraformaldehyde
PKC	Fast-acting protein kinase C
PSC	Pluripotent stem cell
ROCKi	Rho-associated kinase inhibitor
RT-qPCR	Quentative reverse transcription polymerase chain reaction
RT-PCR	Reverse transcription polymerase chain reaction

SB	Sleeping beauty
SOX2	Sex determining region Y-box 2
SOX17	Sex determining region Y-box 17
STAT3	Signal transducer and activator of transcription 3
SUSD2	Sushi containing domain 2
rtTA	Reverse tetracycline transactivator
TAE	Tris/Acetate/EDTA
TBE	Tris/Borate/EDTA
TFCP2L1	Transcription factor CP2 like 1
TGF-β	Transforming growth factor-β
TMP	Trimethoprim
VP16	*Herpes simplex* Viral protein 16
XIST	X-inactive specific transcript

1.Introduction

1.1 Pluripotent stem cells

1.1.1 Naïve and primed states of pluripotency

Stem cells can be characterized based on their two properties: self-renewal and potency (Romito and Cobellis 2016). While all stem cells have unlimited self-renewal, the adult stem cells are considered multipotent which means that they are only capable of differentiating into a specific cell type (Sobhani et al. 2017). Pluripotent stem cells (PSC) on the other hand are able to differentiate into specialized cell types from all of the three (endoderm, ectoderm, mesoderm) germ layers (Wobus and Boheler 2005). PSCs are not found in adults but are present in early embryos (Sobhani et al. 2017). PSCs can be derived from the inner cell mass (ICM) of pre-implantation embryos (Evans and Kaufman 1981; Thomson et al. 1998) or by reprogramming somatic cells by overexpressing specific transcription factors (Takahashi and Yamanaka 2006). Both these embryo derived embryonic stem cells (ESC) and reprogrammed induced pluripotent stem cells (iPSC) can be cultured and expanded *in vitro* indefinitely while maintaining their pluripotent state and the ability to differentiate into all three germ layers (Romito and Cobellis 2016). Since ESCs are derived from embryos they pose ethical concerns as well as possible immunological rejections, so with the use of iPSCs, which can be reprogrammed from the patient's own somatic cells, can overcome these concerns (Romito and Cobellis 2016). Furthermore, in disease modelling a given disorder can be equally modeled with ESCs and iPSCs, but the use of ESCs would require genetic modifications to induce the specific mutation while the patient derived iPSCs would already contain them (Halevy and Urbach 2014).

PSCs exists in different states during the embryo development. The most well studied states, the primed and naïve pluripotent states, can be characterized based on their morphology, gene expression, epigenetic state and ability to contribute to the formation of blastocyst chimeras (Nichols and Smith 2009). The naïve state resembles the ICM of the preimplantation embryo, while the primed state resembles the single cell layer of the postimplantation epiblast found later during the embryo development (Nichols and Smith 2009). The morphology of naïve PSCs is a round, domed cell colony (Weinberger et al. 2016) while the primed PSCs grow as a flat monolayer (Tesar et al. 2007). Primed PSCs represent a later stage in embryo development and depending on the culture conditions or the inductive history of the embryo they can be biased on their ability to

differentiate into different lineages (Nichols and Smith 2009). Primed and naïve PSCs both express the core pluripotency factors such as octamer-binding transcription factor 4 (OCT4), sex determining region Y-box 2 (SOX2) and NANOG but they differ in the expression of other transcripts (Nichols and Smith 2009). Primed state PSCs also cannot contribute to blastocyst chimeras unlike naïve PSCs (Guo et al. 2009; Tesar et al. 2007).

Mouse ESCs are conventionally derived and cultured in the naïve state (Boroviak et al. 2014, 2015), but can also be derived in the primed state as epiblast stem cells (EpiSC) from the postimplantation epiblast (Brons et al. 2007; Tesar et al. 2007). Human derived ESCs are more similar to the mouse EpiSCs than mouse ESCs and also represent the primed state of pluripotency (Davidson et al. 2015; Nakamura et al. 2016). However, the naïve state of pluripotency observed in rodent embryos is also present at primate embryos in the preimplantation epiblast with some species-specific features (Boroviak et al. 2015; Nakamura et al. 2016; Roode et al. 2012; Takashima et al. 2014). Female primed PSCs also show X-chromosome inactivation (Guo et al. 2009, Bao et al. 2009) while female naïve PSCs have two active X-chromosomes in both humans and mice (Barakat et al. 2015). The X-inactive specific transcript (XIST) and histone protein H3 lysine 27 tri-methylation (H3K27me3) both become enriched in primed PSCs containing the inactivated X-chromosome (Guo et al. 2009; Maherali et al. 2007) which can potentially be utilized as means of characterization between the female naïve and primed PSCs.

1.1.2 Signaling pathways

The ICM of a preimplantation blastocyst, also known as naïve epiblast, is unique to the mammals (Nichols and Smith 2009). Mouse naïve stem cells were first successfully derived by coculturing them with inactivated mouse embryo fibroblasts (iMEF) which produce cytokine leukemia inhibitory factor (LIF) (Nichols and Smith 2009). By activating signal transducer and activator of transcription 3 (STAT3), LIF inhibits differentiation and improves naïve PSC viability (Smith 2001). In addition to LIF naïve PSCs require for self-renewal serum or bone morphogenetic proteins (BMP), which are members of the transforming growth factor-β (TGF-β) family (Ying et al. 2008). Primed state PSCs require both fibroblast growth factor (FGF) and TGF-β signaling for maintaining their pluripotency (Vallier et al. 2005). However, FGF also stimulates the mitogen-activated protein kinase (MAPK/ERK) pathway, which in turn primes naïve PSCs for differentiation (Burdon et al. 1999; Kunath et al. 2007; Stavridis et al. 2007). MAPK/ERK pathway can also be activated by other endogenous factors and serum components (Nichols and Smith 2009). Since neither LIF nor BMP blocks the MAPK/ERK pathway activation, naïve PSCs start to prime because of FGF secreted by the cells, unless small-molecule inhibitors are added to the

culture medium (Ying et al. 2003; Kunath et al. 2007). The addition of PD0325901, a highly selective non-competitive inhibitor of the MAPK/ERK kinase (MEK), inhibits FGF receptor tyrosine kinases and the MAPK/ERK cascade (Ying et al. 2008). Therefore, MEK inhibition supports the self-renewal of naïve PSCs by inhibiting the priming of the cells and replacing the need for BMPs. However, without LIF the naïve PSCs will start to degenerate even if the priming is inhibited by PD0325901, so LIF is still required in the culture medium (Ying et al. 2008).

Wnt/ β-catenin is a key pathway for PSCs during the embryo development (Tanaka et al. 2011; Valenta et al. 2012; Nusse and Varmus 2012) and its signaling has been shown to synergize with LIF induced JAK/STAT3 signaling, thus promoting the self-renewal of naïve PSCs and inhibiting differentiation (Hao et al. 2006, Ogawa et al. 2006). Paracrine and autocrine Wnt signaling also inhibits naïve PSCs from priming (ten Berge et al. 2011). Wnt/ β-catenin signaling can be improved by inhibiting glycogen synthase kinase-3 (GSK3), which is involved in the β-catenin destruction complex together with AXIN, adenomatous polyposis coli protein and casein kinase 1α (Xu et al. 2016). β-catenin represses TCF3 (Wray et al. 2011) which is a transcriptional repressor for pluripotency genes such as *NANOG* and transcription factor CP2 like 1 (*TFCP2L1*) (Martello et al. 2012). The use of GSK3 inhibitor CHIR99021 enhances naïve PSCs survival but also promotes non-neural differentiation (Ying et al. 2008). However, the use of CHIR99021 in conjunction with PD0325901 is sufficient to maintain full pluripotency and expand stable naïve PSC colonies in serum-free medium, which has been done successfully with multiple mouse strains (Silva and Smith 2008; Ying et al. 2008). This culture medium with dual inhibition of MEK and GSK3 kinases is also called the 2i medium (Silva et al. 2008). Furthermore, the 2i medium can also be supplemented with a fast-acting protein kinase C (PKC) inhibitor Gö6983, which has been shown to improve maintaining the naïve state of PSCs (Dutta et al. 2011).

1.2 Naïve PSC reversion

1.2.1 Chemical resetting by HDACi

Human primed PSCs can be reprogrammed back to the naïve state by transfection of naïve factors like KLF4 (Guo et al. 2009) or by chemical resetting with transient histone deacetylase (HDAC) inhibition (Guo et al. 2017; Ware et al. 2009). Since human derived ESCs are conventionally in the primed state this provides a way to generate human naïve PSCs *in vitro* from either human ESCs or iPSCs, which can then be further maintained and expanded using the 2i medium. The chemically reset PSCs induced in 2i medium activate the expression of naïve PSC markers, and therefore their transcriptome is more similar to the human ICM than to the primed PSCs (Guo et al. 2017). The

naïve PSC expansion medium used during the chemical reversion protocol, t2iLGö, consists of essential 6 basal medium supplemented with LIF and previously described small-molecule inhibitors: MEK inhibitor PD0325901, GSK3 inhibitor CHIR99021 and PKC inhibitor Gö6983. HDAC inhibitors (HDACi), such as valproic acid and sodium butyrate (NaB), are epigenetic destabilizers that have been used for chemical resetting of mouse primed EpiSCs (Ware et al., 2009) and for human primed PSCs (Guo et al. 2017). In the reset protocol primed PSCs are initially reset with transient HDAC inhibition for three days, and then later expanded in t2iLGö medium (Guo et al. 2017).

The chemical resetting of primed PSCs is a gradual process, that can take from six to nine weeks to complete. The reset PSCs start to form morphologically domed colonies that still express pluripotency markers such as OCT4 and NANOG, but the expression is decreased without the initial HDACi treatment along with increased differentiation (Guo et al. 2017). With the HDACi treatment the reset PSCs expression of OCT4 remains on a similar level to primed PSCs while NANOG expression increases and a primate naïve marker, Krüppel-like factor 17 (KLF17), which is not expressed in primed human ESCs (Guo et al. 2016) becomes detectable within the first week of reversion (Guo et al. 2017). Reset PSCs also start to express naïve transcription factor TFCP2L1, which is expressed in ICM but not in the primed PSCs (Takashima et al. 2014). HDACi seems to promote resetting by making the chromatin more open, which reduces the silencing of naïve transcription factors (Guo et al. 2017). The reset PSCs also undergo metabolic reprogramming during the reversion, from glycolytic metabolism in primed PSCs to oxidative phosphorylation in naïve PSCs (Takashima et al. 2014; Zhou et al. 2012), with genes associated with mitochondria and oxidative phosphorylation being upregulated in the reset PSCs (Guo et al. 2017). The upregulated genes in reset PSCs are also expressed in the ICM of humans, while in these genes are absent or have lower expression in the primed PSCs (Guo et al. 2017). Therefore, the transcription factors encoded by the genes can be utilized as potential markers for naïve pluripotent stem cells (Guo et al. 2017).

Similar to the human ICM, the reset PSCs also have a reduced CpG methylation compared to the high (85-90%) DNA methylation in primed PSCs, although this loss of methylation during the resetting is not uniform in the whole genome of the reset PSCs (Guo et al. 2017). DNA methyltransferase 3 like (DNMT3L) is expressed in reset human naïve cells, which could facilitate *de novo* methylation (Guo et al. 2017). Also, during the reversion silenced X-chromosome present in female primed PSCs is reactivated (Guo et al. 2017). The reset PSCs maintain full pluripotency and can be later re-primed by simply transferring the cells into a matrigel or geltrex coated plates and culturing them with the same essential 8 (E8) medium that is used in primed PSC cultures (Guo

et al. 2017). These re-primed PSCs are also able to successfully differentiate into all three germ layers in embryoid body (EB) assay, but the naïve PSCs cultured in t2iLGö are not responsive to lineage differentiation and require a re-priming first in order to form EBs (Guo et al. 2017).

1.2.2 Improving the stability of reset PSCs

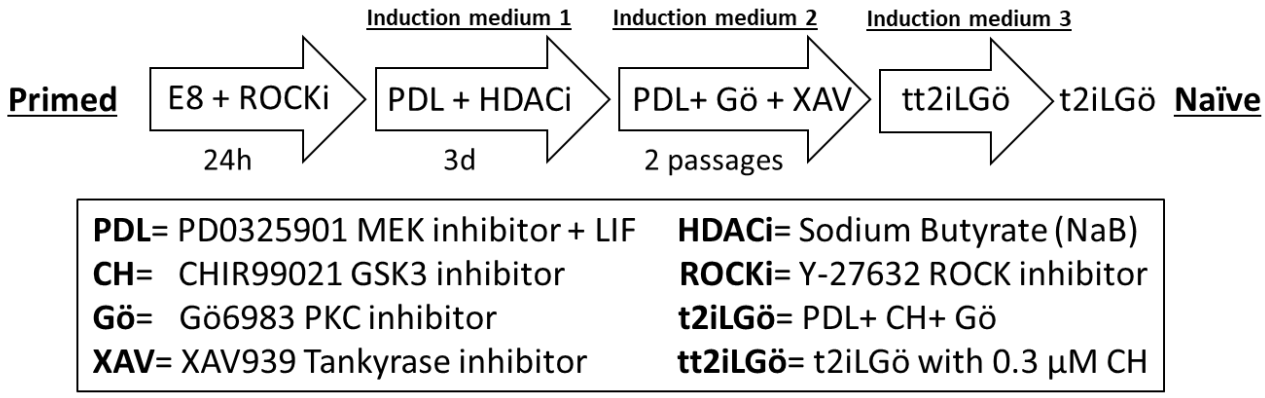


Figure 1. The stabilized naïve reversion protocol with initial HDACi resetting using three induction mediums

After changing the medium into t2iLGö the reset PSCs become stable after few passages, with the resetting efficiency varying between cell lines and the condition of the cell culture prior to resetting (Guo et al. 2017). The reversion efficiency and the stability of reset PSCs can be improved by using multiple induction mediums supplemented with different small-molecule inhibitors during the reversion protocol, before transferring the cells into t2iLGö medium for expansion (Figure 1). Rho-associated kinase inhibitor (ROCKi) Y-27632 is commonly used with human ESCs to improve cell survival by decreasing the dissociation-induced apoptosis during passaging (Watanabe et al. 2007). ROCKi Y-27632 can also be utilized during the reversion protocol in order to improve cell survival during cell passaging (Guo et al. 2017). Reset PSCs are cultured on an iMEF feeder layer but can also be cultured under feeder-free conditions (Takashima et al. 2014). However, the amount of differentiation and cell death is increased in the cell cultures without the use of feeder layer and the reset phenotype does not sustain without the use of feeders (Guo et al. 2017). Without feeders some reset cell lines expand better in low CHIR99021 concentrations or even without it, which has also been observed previously using the alternative 5i naïve culture where GSK3 inhibition is considered optional (Theunissen et al. 2016). Omitting CHIR99021 at the start of the reversion protocol improves the efficiency of initial resetting of the cells, while the morphology and naïve marker expressions remains the same (Guo et al. 2017). Naïve cells cultured in 5i medium are also prone to

aneuploidy (Pastor et al. 2016; Sahakyan et al. 2017; Theunissen et al. 2014) and the manipulation of culture conditions increases the risk of producing genetic variants (Amps et al. 2011). In t2iLGö medium naïve PSCs can maintain their diploid karyotype, although abnormalities may occur after extended (40+ passages) cell culturing (Guo et al. 2016; Takashima et al. 2014).

In addition to omitting CHIR99021 at the start of the protocol, Wnt inhibition can be used to stabilize the resetting process (Guo et al. 2017). Tankyrase inhibitor XAV939, which inhibits canonical Wnt signaling by stabilizing AXIN in the β-catenin destruction complex (Huang et al. 2009), facilitates propagation of PSCs during alternative states and has been shown to improve stability of naïve PSCs in 2i medium (Kim et al. 2013; Zimmerlin et al. 2016). The mechanism of how XAV939 stabilizes naïve PSCs is unclear, but since canonical Wnt signaling is determined by the balance of activated β-catenin and its destruction complex, the synergy between XAV939 and CHIR99021 may stabilize the AXIN-catenin complex, thus increasing cytoplasmic β-catenin retention (Kim et al. 2013; Schmitz et al. 2013) and also improving the stability of reset cells during the reversion. Addition of XAV939 after the initial resetting phase with HDACi results in morphologically more domed and stable reset PSC colonies, with reduction in differentiation and cell death (Guo et al. 2017). These reset PSCs treated with XAV939 also express higher levels of naïve markers compared to the untreated PSCs and can be expanded with or without further Wnt inhibition (Guo et al. 2017). After two passages in medium supplemented with XAV939, the reset PSCs can then be transferred into t2iLGö medium with CHIR99021 concertation lowered from 1 µM to 0.3 µM. This medium with a lowered CHIR99021 concertation is also called tt2iLGö (Guo et al. 2017). Not all of the cells are reset during the initial resetting phase with HDACi and will instead remain as background cells resembling primed or differentiated morphology during the reversion protocol. XAV939 may also improve the stability of these background cells since they will start to rapidly disappear from the culture after the change into tt2iLGö medium. After all of the background cells have disappeared, the finished reset PSCs can be transferred and expanded in the t2iLGö medium, where they can be characterized as naïve PSCs based on their morphology and expression of naïve markers (Guo et al. 2017).

1.3 Naïve markers

While both primed and naïve human PSCs express the core pluripotency factors, they do differ in the expression of other transcripts (Nichols et al. 2009). In addition to the difference in morphology between the primed and naïve PSCs, these multiple genes that are expressed in naïve PSCs but not in primed PSCs can be utilized for characterization purposes using immunocytochemistry and

quantitative reverse transcription polymerase chain reaction (RT-qPCR) assay. In the case of female PSC lines, the H3K27me3 trimethylation marking the silenced X-chromosome can also be utilized for characterization of primed PSCs (Guo et al. 2009; Maherali et al. 2007).

1.3.1 NANOG

Together with OCT4, NANOG is a transcriptional regulator specifically associated with the establishment and maintenance of pluripotency (Nichols and Smith 2009). OCT4 is expressed broadly in PSCs, but NANOG expression is more correlated with the ICM (Chambers et al. 2003). NANOG appears in random cellular distribution during compaction, and it is not inherited maternally (Dietrich and Hiiragi 2007). During the development of blastocyst NANOG is confined to the ICM, and after the implantation the expression of NANOG is downregulated (Chambers et al. 2003). The expression of NANOG is not absolutely necessary for PSCs, but the absence of NANOG is known to increase differentiation (Chambers et al. 2007). Therefore, it is possible that NANOG is more involved in establishing the naïve state and has a secondary role in the self-renewal of the naïve PSCs (Silva and Smith 2008). NANOG expression is present in both primed and naïve PSCs, but the expression level is increased during the primed to naïve reversion (Guo et al. 2017).

1.3.2 KLF17

KLF17 is a naïve marker exclusively observed in the primate ICM replacing other factors found in mouse naïve PSCs, such as KLF2, ESRRB and BMP4 (Blakeley et al. 2015; Boroviak et al. 2015). KLF17 is initially expressed at the 8-cell stage and later enriched in the ICM, co-localizing with NANOG (Blakeley et al. 2015). When naïve PSCs are primed, their NANOG expression decreases and the KLF17 expression disappears (Guo et al. 2016). There is little functional data on KLF17, but it is known to act as a negative regulator for the epithelial-mesenchymal transition and metastasis in breast cancer (Gumireddy et al. 2009).

1.3.3 TFCP2L1

TFCP2L1 is one of the key targets in LIF/STAT3 mediated self-renewal in naïve PSCs (Ye et al. 2014). TFCP2L1 upregulates NANOG expression and thus promotes the self-renewal of naïve PSCs in a NANOG-dependent manner (Ye et al. 2014). TFCP2L1 is highly upregulated in the human ICM but becomes downregulated in the primed PSCs (O'Leary et al. 2012). Forced expression of TFCP2L1 and NANOG can also replace the need for MEK inhibition by PD0325901 in maintaining the naïve state, but neither is essential for maintaining the naïve state when using the

2i medium (Ye et al. 2014). Forced expression of TFCP2L1 can also be used to reprogram primed PSCs back to the naïve state (Ye et al. 2014).

1.3.4 H3K27me3 and H9 cells

The inactivation of the X-chromosome in female primed PSCs is an epigenetic change between naïve and primed PSCs with the enrichment of XIST and H3K27me3 marker expressions (Guo et al. 2009; Maherali et al. 2007). H3K27me3 methylation is deposited at CpG-dense promoters by Polycomb PRC2 complex and it maintains gene transcriptional repression (Morey and Helin 2010). Female primed PSCs show a genome-wide increase in the H3K27me3 marker (Yang et al. 2017) while in naïve PSCs the H3K27me3 marker is almost entirely lost (Theunissen et al. 2014). Observing the H3K27me3 methylation can be utilized for the characterization of female primed PSC lines, such as H9. H9 is a female ESC line derived in the University of Wisconsin (Thomson et al. 1998) and is alongside with the H1 cell line one of the most researched and distributed human ESC lines (Chen et al. 2008). H9 cells have also been previously used successfully in the chemical HDACi naïve reversions (Guo et al. 2017).

1.3.5 DNMT3L

DNMT3L is a catalytically inactive DNA methyltransferase that is highly expressed in naïve PSCs (Neri et al. 2013). DNMT3L positively regulates methylation at gene bodies of housekeeping genes, and negatively at bivalent genes promoters (Neri et al. 2013). By interacting with Polycomb PRC2 complex, DNMT3L competes with other methyltransferases such as DNMT3a and DNMT3b to maintain low methylation at H3K27me3 regions (Neri et al. 2013). Therefore, DNMT3L is counteracting other *de novo* DNA methylases in order to maintain hypomethylation of bivalent genes promoters (Neri et al. 2013).

1.4 Endogenous gene expression by CRISPRa

1.4.1 CRISPR/Cas9

Bacteria and archaea can target foreign DNA using the clustered regularly interspaced short palindromic repeat (CRISPR) system (Mali et al. 2013). Using sequence-specific guide RNAs (gRNA) CRISPR associated proteins (Cas) can be targeted to complementary sites of foreign viral or plasmic DNA (Jinek et al. 2012). CRISPR system derived from *Streptococcus pyogenes* is classified as a type II CRISPR system which consists of Cas9 protein and a gRNA (Jinek et al. 2012). Cas9 has homologous domains to HNH and RuvC endonucleases which both cut one strand

of the DNA causing a double stranded break (Jinek et al. 2012). The gRNA consists of a specific target sequence that recognizes crRNA, which binds to Cas9 via complementary tracrRNA (Jinek et al. 2012). These RNAs can also be fused together by a hairpin loop into a single chimeric gRNA (Jinek et al. 2012). Type II CRISPR/Cas9 systems ability to cause mutations via double-stranded break at the targeted sites can be used as a gene editing tool even in eukaryotic cells (Jinek et al. 2012; Nishimasu et al. 2015). gRNAs are expressed under Pol III promoters like U6, and for *in-vitro* experiments, multiple gRNA cassettes containing the U6 promoter and the sequence-specific gRNA can be assembled into a single plasmid vector using Golden Gate cloning (Vad-Nielsen et al. 2016; Weltner 2018). In the combined vector each gRNA is regulated by its own U6 promoter (Vad-Nielsen et al. 2016). The use of multiple gRNAs can improve the efficiency of Cas9 binding into the target site or allow the targeting of multiple genes by using gRNAs that target different sites (Vad-Nielsen et al. 2016).

1.4.2 CRISPRa

Cas9 can also be catalytically inactivated so that the protein is unable to cut the DNA double-strand at the target site (Jinek et al. 2012). Instead the catalytically inactivated "dead" Cas9 (dCas9) can be fused with repeats of the *Herpes simplex* viral protein 16 (VP16) transactivator domain to alter gene expression at the target site (Maeder et al. 2013; Perez-Pinera et al. 2013). When targeted upstream of transcription start site of the target gene, dCas9VP16 can recruit RNA polymerase II to the promoter site thus activating the transcription of the target gene (Perez-Pinera et al. 2013; Weltner 2018). Target gene activation can be further improved by increasing the number of VP16 repeats, for example with 12 repeats titled VP192 (Balboa et al. 2015; Cheng et al. 2013; Maeder et al. 2013; Mali et al. 2013; Perez-Pinera et al. 2013). Using this CRISPR mediated activation (CRISPRa) we can utilize the dCas9VP192 fusion transactivator in order to activate endogenous expression of targeted genes and further study their effects on naïve PSCs and during the naïve reversion.

The activity of the dCas9VP192 transactivator can be controlled by fusing it with either a doxycycline (DOX) inducible TetON-promoter or with a dihydrofolate reductase (DHFR) destabilization domain (DD) (Balboa et al. 2015). In TetON system the tetracycline response element is located upstream of the dCas9s CMV promoter region, and the reverse tetracycline transactivator (rtTA), which contains a VP16 activation domain, is only able to bind to the promoter region at the presence of DOX (Gossen and Bujard 1992). When fusing the *E. Coli* derived DHFR DD with the dCas9, the protein becomes unstable which causes protein degradation in the absence of trimethoprim (TMP) that inhibits DHFRs ability to catalyze the reduction of dihydrofolate to

tetrahydrofolate (Iwamoto et al. 2010). Even when the gRNA of target gene is present, both the TetON and DHFR DD systems are capable of reducing the unwanted leaky expression of the target gene when the cells are not induced by DOX or TMP (Balboa et al. 2015). However, the combination of both TetON and DHFR DD systems can further eliminate the leaky activation of target genes when DOX and TMP are not present, so using both of these systems together allows the generation of conditionally stabilized DDdCas9VP192 activator system, where the expression of target genes is activated only when the cells are induced with both DOX and TMP (Balboa et al. 2015; Weltner et al. 2018).

1.4.3 Transfection by DNA transposons

DDdCas9VP192 can be transfected into cells via plasmid vectors. Non-integrative episomal plasmids containing the Epstein-Barr virus origin of plasmid replication and EBNA-1 sequences can be used in somatic cell reprogramming, however these plasmids are transiently lost during proliferation (Hu 2014; Weltner 2018). In order to create a sustainable DDdCas9VP192 activator cell line, which can then be used for studying the effects of selected genes in naïve PSCs and during the reversion, DNA transposons can be used instead to mediate the plasmid sequence's integration to the host genome. Sleeping beauty (SB) is the first transposon system used in mammals and a member of Tc1/mariner superfamily (Ivics et al. 1997; Wang et al. 2008). SB is an ancestral fish gene that was engineered by eliminating the mutations that were inactivating the gene (Ivics et al. 1997). Another transposon system is based on piggyBac (PB) transposon which was isolated from *Trichoplusia ni* (Cary et al. 1989) and has a higher activity in mammals compared to the SB (Wang et al. 2008).

Both transposon systems work via a "cut and paste" mechanism, where the transposase recognizes the terminal inverted repeat sites at both sides of the transposon vector, moving it from its original location to a new location in the genome based on the sequence targeted by the transposase (Ivics et al. 1997). SB transposase insert the SB vector into a sequence containing TA nucleotides (Plasterk et al. 1999), while PB targets TTAA sites (Fraser et al. 1983). In order to integrate the sequence of interest, for example the DDdCas9VP192 sequence, to the target genome it must have the terminal inverted repeat sites of the selected transposon on both ends of the sequence. A second helper plasmid expressing the selected transposase is transfected at the same time, integrating the sequence into the target genome (Weltner 2018; Wang et al. 2008). The transfected cells can then later be selected by green fluorescence protein (GFP) expression or by antibiotic selection, provided that the transfected sequence also contained a GFP or an antibiotic resistance sequence downstream from the promoter region.

1.4.4 LEUTX

Leucine twenty homeobox (*LEUTX*) is a PRD-like homeobox gene which in humans is expressed exclusively in the early embryos (Jouhilahti et al. 2016). *LEUTX* is characterized by the PRD class homeodomain with the exception of leucine at position 20, and like other PRD-like transcription factors it lacks the paired domain (Galliot et al. 1999; Bürglin 2011). *LEUTX* is almost exclusive to humans, with related sequences detected in primates and divergent orthologs in other placental mammals (Zhong et al. 2011). LEUTX is expressed during preimplantation embryos peaking at the 4-cell stage but is not expressed later in the adult cells, and when transiently expressed in human PSCs, LEUTX is capable of activating the transcription of multiple genes upregulated during the human embryonic genome activation (EGA), including pluripotency-associated genes (Jouhilahti et al. 2016).

LEUTX is suggested to play an important role during preimplantation development and to act as a main regulator of EGA alongside DUXA and OTX2 (Töhönen et al. 2015) by activating multiple pluripotency and EGA-associated genes (Jouhilahti et al. 2016). These genes are later repressed by DPRX, which competes on the same 36bp DNA motif with LEUTX, countering the effects of LEUTX (Jouhilahti et al. 2016). The effect of LEUTX on naïve PSCs is still unknown. The involvement of LEUTX during EGA and in the activation of multiple pluripotency and EGA-associated genes raises the question whether the expression of LEUTX has an effect on naïve reversion. LEUTX expression could potentially improve the reversion efficiency or the maintenance of naïve PSCs by driving the transcriptome of primed PSCs back towards the earlier cell stages of embryo development, potentially even to the cell stages preceding the naïve state. Using the DDdCas9VP192 transactivator system we can activate the endogenous expression of LEUTX prior to the initial resetting phase in the naïve reversion and study its effects during the reversion process.

2. Aims of the thesis

The chemical reversion of primed to naïve PSCs via transient HDAC inhibition is a recently developed method that allows the reversion of primed human PSCs back to the earlier naïve state. These naïve PSCs can be utilized for further studying and modelling of earlier cell stages in human embryo development. Furthermore, a CRISPRa-based conditionally stabilized DDdCas9VP192 activator cell line allows the sequence-specific activation of endogenous genes in order to study their effects on the PSCs during the reversion protocol. Since *LEUTX* plays an important role during preimplantation development and in pluripotency, we selected it as the target gene for the reversion. The expression of LEUTX can activate multiple early embryo genes, which could improve the naïve reversions efficiency, as well as the maintenance and proliferation of the generated naïve PSC colonies.

The aims of this thesis were to:

- Set up and optimize the HDACi based reversion protocol in order to revert primed H9 cells back to the naïve state
- Characterize the reverted cells based on their morphology and selected naïve marker expressions using immunocytochemistry and RT-qPCR assays.
- Generate a functional DDdCas9VP192 activator cell line using H9 PSCs to study the effect of the endogenous expression of LEUTX during the reversion.

3. Materials and methods

3.1 Cell culturing and imaging

Primed human H9 ESCs (WA09, WiCell) were cultured in pluripotent stem cell medium E8 (Gibco) at 37 °C, 20 % O_2, 5 % CO_2 growth conditions in a Forma Steri-Cycle CO2 incubator (Thermo Fisher). Cell culture plates were coated with Matrigel (Corning) or Geltrex (Gibco). Naïve H9 cells were cultured in t2iLGö Naïvecult expansion medium (Stemcell Technologies) on iMEF (Gibco) feeder layer of 3 x 10^5 cells at hypoxic 37 °C, 5 % O_2, 5 % CO_2 growth conditions in X-vivo workstation (BioSpherix). The medium was changed daily for the naïve PSCs, and primed PSCs were cultured up to 48 hours in E8 medium or up to 72 hours in E8 Flex medium (Gibco) before the medium change. All cell work was done in Biowizard silver line biosafety cabinets (Kojair Tech). For light microscopy a Leica DM IL LED microscope (Leica-microsystems) was used, and pictures of live cells were taken with an attached Leica EC3 (Leica-microsystems) camera using LAS-EZ imaging software (Leica-microsystems). EVOS FL cell imaging system (Thermo Fisher) was used for fluorescence imaging. During the naïve reversions whole culture plate-images for colony counting were taken with the IncuCyte S3 live-cell analysis system (Sartorius).

When passaging the primed PSCs, old medium was aspirated from the culture plates and the cells were washed twice with 0.5 mM ethylenediaminetetraacetic acid (EDTA, Life Technologies) and incubated with EDTA for 3-4 minutes at 37 °C. After the incubation EDTA was aspirated, and the cells were detached by pipetting them carefully with new medium 2-3 times or by scraping the cells. The detached cells were then transferred and split into new plates in preferred ratio. Single cell passaging was used for passaging naïve PSCs and during electroporations. After the medium aspiration TrypLE (Life Technologies) was added to the cells, and then incubated for 3 minutes at 37 °C. After the incubation PBS supplemented with 10 % fetal bovine serum (FBS) was added to neutralize the passaging reagent. The whole solution was then moved into 15 ml falcon centrifuge tubes and centrifuged in an Eppendorf 5810R centrifuge (Sigma) at 300 g for 5 minutes. After centrifugation the supernatant was aspirated, and the cell pellet resuspended in new medium. The cells were counted by adding 10 µl of trypan blue dye to 10 µl of the cell suspension and then counted with Countess Cell counter (Invitrogen). The cell medium was also supplemented with 10 µM ROCKi Y-27632 (Wako Chemicals) for 24 hours after single cell passaging.

3.2 Cryopreservation and thawing of cells

The cells were passaged either by EDTA or single cell passaging reagents, but instead of adding new medium, they were first suspended in 0.22 μm filter sterilized freezing Medium 1 consisting of DMEM (Sigma) with 20 % FBS. The suspended cells were then transferred into a cryo vial, and an equal amount of 0.22 μm filter sterilized freezing medium 2 was added. Freezing medium 2 consists of DMEM with 20 % FBS and 20 % of dimethyl sulphoxide (DMSO) (Sigma). The vials were first frozen in polystyrene boxes at -80 °C freezer overnight, and later transferred into -150 °C freezer for long term storage.

When thawing cells, the cryo vial was removed from the -150 °C freezer and thawed in either a water bath or by hand, until only a small crystal of ice remained. The vial was wiped with 70 % ethanol before pipetting the cells into a 15 ml falcon centrifuge tube containing 4 ml of medium. Cells were then centrifuged at 300 g for 5 minutes, and after the supernatant was removed the cell pellet was resuspended in a new medium, supplemented with 10 μM ROCKi Y-27632. The suspended cells were then moved into new culture plates coated with Matrigel. The iMEF feeder layers were thawed the same way but on 0.1 % gelatin covered plates and without the addition of ROCKi Y-27632.

3.3 RT-PCR

Cells were lysed in 350 μl of LBP lysis buffer included in the NucleoSpin RNA Plus kit (Macherey-Nagel), which was used for the isolation of the RNA. RNA concentrations were measured with Simplinano spectrophotometer. Complementary DNA (cDNA) was synthesized from 1 μg of RNA by reverse transcription. Molecular biology grade water (Lonza) was used for both the reverse transcription polymerase chain reaction (RT-PCR) and the RT-qPCR reactions, as well as for all the other assays described in materials and methods.

RT reaction mix	**V(μl)**
M-MLV Reverse transcriptase buffer 5x (Promega)	4
dNTP 2.5 mM (Promega)	2.5
Oligo dT primer (Thermo Scientific)	1
Random hexamers (Promega)	0.2
RNAse inhibitor (Thermo Scientific)	0.5
M-MLV Reverse transcriptase (Promega)	0.5
H_2O + 1 μg RNA	11.3

H_2O was added to the 1 µg of RNA for a total volume of 11.3 µl, and the RNA was denatured in 65 °C for 1 minute. 8.7 µl of reverse transcription master mix containing reverse transcriptase was prepared and added per reaction. After the master mix was added to the RNA and H_2O, the whole mix was spun down and incubated at 37 °C for 90 minutes. After incubation the reverse transcriptase was inactivated at 95 °C for 5 minutes. For the RT-PCR reaction 1 µl of synthesized cDNA was used per reaction. cDNA was added to 19 µl of master mix containing DreamTaq DNA polymerase (Thermo Scientific). Primers used in RT-PCR and RT-qPCR assays are listed in table 1.

PCR reaction mix	**V(µl)**
DreamTaq Green Buffer 10x (Thermo Scientific)	2
dNTP 2.5mM (Promega)	1.6
Forward Primer 10 µM	0.5
Reverse Primer 10 µM	0.5
DreamTaq DNA polymerase (Thermo Scientific)	0.1
H_2O	14.3
cDNA	1

The PCR reactions were done with S1000 Thermal Cycler (Bio-Rad) for 35 cycles. A negative water control was always used in addition to the samples.

PCR reaction	
98°C Denaturation	3 min
35 cycles	
98°C Denaturation	10 sec
Annealing temperature	30 sec
72°C Elongation	12 sec
72°C Elongation	8 min

After the PCR reaction was complete, the products were run by electrophoresis on 100 ml 2.5 % Tris/Borate/EDTA (TBE) agarose-gel supplemented with 5 µl of Midori green advance DNA stain (Nippon Genetics). 10 µl of PCR product and 5 µl of GeneRuler 1kb Plus DNA-ladder (Thermo Scientific) was loaded on the gel, and the electrophoresis ran at 120 V for 30 minutes. The bands could be then observed with Bio-Rad Gel doc XR+ gel documentation system.

Table 1. RT-PCR and RT-qPCR primers

Gene	Forward Primer	Reverse Primer	Annealing T (°C)	Amplicon (bp)
NANOG	5’ CTC AGC CTC CAG CAG ATG C	5’ TAG ATT TCA TTC TCT GGT TCT GG	57	94
KLF17	5’ CTG CAA CTA CGA GAA CTG CG	GCA AGA ATA TGG CCT CTC AC	60	94
KLF17 (2)	5’ CCG CAG GCT GAG ATG GAA CA	5’ ATG GGC GCT GAG TTC TCG TT	60	92
DNMT3L	5’ CTG CTC CAT CTG CTG CTC C	5’ ATC CAC ACA CTC GAA GCA GT	60	85
DNMT3L (2)	5’ TGG AAA GTG GTT CTG ACC CG	5’ AAG ATC GAA GGG TCC CCA CT	60	95
TFCP2L1	5’ TTT GTG GGA CCC TGC GAA G	5’ TGC TTA AAC GTG TCA ATC TGG A	60	129
LEUTX	5’ GCT ACA ATG GGG AAA CTG GC	5’ CTC TTC CAT TTG GCA CGC TG	60	89
Cyclophilin G	5’ TCT TGT CAA TGC CCA ACA GAG	5’ GCC CAT CTA AAT GAG GAG TTG	60	84
dCas9	5’ TGA TCG AGA CAA ACG GCG AA	5’ CTG TCT GCA CCT CGG TCT TT	60	1934
SOX17	5’ CCG AGT TGA GCA AGA TGC TG	5’ TGC ATG TGC TGC ACG CGC A	57	103
FOXA2	5’ AAG ACC TAC AGG CGC AGC T	5’ CAT CTT GTT GGG GCT CTG C	57	93

3.4 RT-qPCR

The RNA was isolated, and cDNA synthesized just as previously described with RT-PCR. 1 µl of cDNA was added to a master mix containing 4 µl of 5x HOT FIREPol EvaGreen qPCR Mix Plus (ROX) (Solis Biodyne) and 10 µl of H_2O per reaction. Qiagen QIAgility pipetting robot was used for pipetting 15 µl of sample mix with 5 µl of 2 µM Forward/Reverse primer mix per reaction. All reactions were done as duplicates. In addition to the samples a negative water control was always used. Cyclophilin G was used as housekeeping gene for normalizing the results. After pipetting the RT-qPCR reaction was done by Qiagen Rotor-Gene Q qPCR machine.

qPCR reaction

95 °C Denaturation	15 min
35 cycles	
95 °C Denaturation	25 sec
Annealing temperature	25 sec
72 °C Elongation	25 sec

At the end of the RT-qPCR reaction sample's melting curve was also analyzed by increasing the temperature starting from 75 °C by 1 °C every 5 seconds. The qPCR results were analyzed by ddCt method, using Cyclophilin G as a housekeeping gene for normalization and the expression data was presented as $\log_2$ fold change between the samples.

3.5 Immunocytochemistry

The cells were plated into 24-well culture plates prior to the immunostainings. Medium was aspirated from the wells and the cells were washed with PBS. After the wash cells were fixed with 250 µl of 4 % paraformaldehyde (PFA) (Fisher Chemical) in PBS for 15 minutes. The PFA was removed and the cells were washed again with PBS. The cells were then permeabilized by 250 µl of 0.5 % Triton X-100 in PBS for 10 minutes. After triton was aspirated, cells were washed with PBS and then 4 drops of UV-blocker (Thermo Scientific) was added on the wells for 10 minutes to block epitopes. UV-blocker was aspirated and 200 µl of primary antibody diluted in 0.1% Tween in PBS was added to the wells (Table 2). Wells were incubated for 24 hours in dark at 4 °C on a Stuart SSL4 see-saw rocker.

Table 2. Primary antibodies

Primary Antibody	**Manufacturer**	**Host Species**	**Dilution**	**Catalog**
NANOG	Cell signaling technologies	Rabbit	1:200	4903S
KLF17	Atlas	Rabbit	1:500	HPA024629
H3K27me3	EMD Millipore	Rabbit	1:500	07-449
LEUTX	Novus Biologicals	Rabbit	1:200	NBP1-90890
FOXA2	Santa Cruz	Goat	1:500	sc-6554
SOX17	R&D systems	Goat	1:500	AF1924
α-Smooth muscle actin	Sigma	Mouse	1:500	A2547
β-tubulin III	Abcam	Rabbit	1:500	Ab18207

After the primary antibody was aspirated wells were washed with PBS before adding 200 µl of secondary antibody and Hoechst 33342 (Thermo Fisher scientific) DNA dye diluted in 0.1 % Tween in PBS (Table 3). The wells were incubated in the dark at RT for 30 minutes on the see-saw rocker. After incubation the secondary antibody was aspirated and the wells washed with PBS, leaving the last PBS on wells to prevent the cells from drying.

Table 3. Secondary antibodies and DNA dye

Secondary Antibody/DNA dye	**Manufacturer**	**Host Species**	**Dilution**	**Catalog**
Hoechst	Thermo Fisher scientific	DNA dye	1:1000	62249
Anti-Rabbit IgG (H+L) Alexa Fluor 594 (Red)	Invitrogen	Donkey	1:500	A21207
Anti-Goat IgG (H+L) Alexa Fluor 594 (Red)	Invitrogen	Donkey	1:500	A11058
Anti-Mouse IgG (H+L) Alexa Fluor 488 (Green)	Invitrogen	Donkey	1:500	A21202
Anti-Goat IgG (H+L) Alexa Fluor 488 (Green)	Invitrogen	Donkey	1:500	A11055

3.6 Transfections

H9 cells were treated with 10 µM ROCKi Y-27632 for four hours before electroporations. Electroporations were done with Neon transfection system (Invitrogen) using the protocol from the Neon transfection system 100 µl kit (Invitrogen). 2 x 10^6 cells were suspended in 200 µl of R-buffer and electroporated using 1100 V, 20 ms, 2 pulses conditions. The electroporated cells were then transferred into a new 10 cm culture plate with added 10 µM ROCKi Y-27632. Activator cell line was generated by electroporating H9 cells with two plasmids containing DDdCas9VP192 (Supplementary Figure 3) and rtTA sequences, which were integrated into the genome by sleeping beauty transposase. Guide plasmids were electroporated into H9 DDdCas9VP192 activator cells and integrated with piggybac transposase (Table 4). The ability of activator cells to activate the targeted genes endogenous expression was tested by electroporating a plasmid containing gRNAs for endoderm transcription factors sex determining region Y-box 17 (SOX17) and forkhead box protein A2 (FOXA2).

Table 4. Transfection plasmids and DNA transposases

Plasmid	**Amount of DNA (ng/1 x 10^6 cells)**	**Plasmid c (ng/µl)**	**V (µl)**
SB-tight-DDdCas9VP192- GFP-Zeo-WPRE	1000	494	2.02
SB-CAG-rtTA-IN-IRES-Neo	1000	548	1.82
CAG-SB-100X-bghpA	500	458	1.09
PB-GG-FOXA2-SOX17-PGK-Puro	1000	305.9	3.27
pSubCMV-hyfBase-WPRE-piggybac	500	82	6.10
PB-GG-LEUTXg4+5-PGK-Puro	1000	443.1	2.23
pCMV-HAhy-PBase	500	1179	0.42

After the transfections, successful activator plasmid integration could be observed in the activator cells by their GFP expression. Since all integrated plasmids contained an antibiotic resistance sequence for either neomycin, zeocin or puromycin, the electroporated cells were selected with antibiotic treatment. For neomycin selections G418 (Life technologies) was used. The optimal concentrations for neomycin and zeocin antibiotic selection were tested by treating H9 cells for four days with different concentrations of antibiotics, and the amount of cell death was observed daily (Supplementary Figure 1). For neomycin different G418 concentrations of 0-500 µg/ml were tested, and for zeocin (Sigma) 0-10 µg/ml. Puromycin (Sigma) selections were done with concentration of 0.5 µg/ml.

3.7 Guide PCR

LEUTX guides were selected from a total of nine different guides, which had been previously tested in the research group. Based on the previous results guides 2-5 were selected for further testing (Table 5).

Table 5. LEUTX guide sequences

LEUTX embryonic promoter guide	**Guide sequence**
Guide 2	GCGTGGTATTAGGGTAGGAC
Guide 3	TATTGGAGGGCGTGGTATTA
Guide 4	GATATTGAATGGATTATTGG
Guide 5	CTGATGCTGTGTAGGGCACT

Guide cassettes were assembled in a PCR reaction using the four different LEUTX guide oligos. 2 µl of guide oligo was added into 98 µl of mastermix.

Guide PCR reaction mix	**V(µl)**
5x HF buffer	20
dNTP 2.5 mM (Promega)	8
1 aggc Forward Primer 100 µM	0.5
1 aggc Reverse Primer 100 µM	0.5
Phusion	1
U6 promoter tailed 5 ng (21 ng/µl)	0.23
Term tailed 5 ng (20.2 ng/µl)	0.25
H_2O	67.52
gRNA oligo (1 µM)	2

Guide PCR	
98 °C Denaturation	3 min
35 cycles	
98 °C Denaturation	10 sec
52 °C Annealing	30 sec
72 °C Elongation	12 sec
72 °C Elongation	8 min

The PCR products were run in 1 % Tris/Acetate/EDTA (TAE) gel, and the products were purified using NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel).
The finished LEUTX guide cassettes were then tested by transfecting them into HEK293 human embryonic kidney cells. HEK293 cells were transfected using FuGENE HD (Promega) protocol. 6 x 10^5 cells were transfected with 500 ng of episomal pCXLE-dCas9-VP192-T2A-GFP plasmid and with 200 ng of each gRNA cassette (Table 6). A negative tdTomato control and a positive pool control containing all of the four guide cassettes were also transfected.

Table 6. Episomal plasmid and guide cassettes for HEK293 transfection

Plasmid/Guide cassette	**Amount of DNA (ng)**	**Concentration (ng/µl)**	**V(µl)**
pCXLE-dCas9-VP192-T2A-GFP	500	1178	0.42
LEUTX g2	200	212.7	0.94
LEUTX g3	200	209.6	0.95
LEUTX g4	200	268.2	0.75
LEUTX g5	200	128.5	1.56
tdTomato	200	993	0.2

3.8 Golden-gate assembly and minipreps

Two guide cassettes were golden gate assembled into a GG-dest plasmid backbone. The guide cassettes were prepared by previously described guide PCR, with one of the guides using 1 aggc forward primer and 1 aggc reverse primer, and the other 2 aggc forward primer and 5 aggc reverse primer (Table 7).

Table 7. Golden gate primers, table edited from Weltner 2018

Primer	**Sequence**	**Compatible with**
1_aggc_Fw	ACTGAATTCGGATCCTCGAGCGTCTC ACCCTG TAAAACGACGGCCAGT	GG-dest plasmid
1_aggc_Rv	CATGCGGCCGCGTCGACAGATCTCGTCTC ACATGA GGAAACAGCTATGACCATG	2_aggc_Fw
2_aggc_Fw	ACTGAATTCGGATCCTCGAGCGTCTC ACATGG TAAAACGACGGCCAGT	1_aggc_Rv
5_aggc_Rv	CATGCGGCCGCGTCGACAGATCTCGTCTC ACGTTA GGAAACAGCTATGACCATG	GG-dest plasmid

Golden Gate reaction mix	**V(µl)**
10x T4 ligase buffer	2
T4 DNA ligase	1
Esp3I	1
DTT 10 mM	2
gRNA 4 cassette 50 ng (172.6 ng/µl)	0.289
gRNA 5 cassette 50 ng (128.5 ng/µl)	0.389
GG-dest plasmid 150 ng (504.1 ng/µl)	0.298
H_2O (total V=20 µl)	13.024

Golden Gate reaction

50 cycles	
37 °C	2 min
16 °C	5 min
80 °C	20 min

The finished golden gate product was transformed and cloned in New England Biolabs 5α transformation competent *E. coli* cells (NEB5), which are a derivative of DH5α *E. coli* cells. 100 µl of NEB5s were thawed and 100 µl of plasmid was added. After incubating the NEB5s for 30 minutes on ice the cells were heat shocked at 42 °C for 1 minute, and then returned to ice for 2 minutes. The cells were then plated into a LB-Ampicillin dish plate and incubated for 24 hours at 37 °C in MIR-154 incubator (Sanyo). Bacteria colonies were then picked from the LB-Ampicillin

dish into a 15 ml falcon centrifuge tube containing 5 ml of LB-Ampicillin and incubated for another 24 hours at 37 °C on a Sanyo orbital shaker. Minipreps of plasmid were then prepared using GeneJET plasmid miniprep kit (Thermo Fisher). The DNA concentrations were measured with Simplinano spectrophotometer.

3.9 Plasmid restriction and ligation

In order to clone the gRNA cassette into PB-GG-PGK-Puro plasmid backbone (Supplementary Figure 4), the insert was first transferred from GG-dest to GG-dest-EBNA backbone using EcoRI and NotI restriction enzymes and further to PB-GG-PGK-Puro backbone using EcoRI and SpeI. The digested GG-dest minipreps were ran on a 1% TAE gel at 120 V for 30 minutes and based on the band quality two of the clones were selected for sequencing. 5 µl of plasmid containing 400-500 ng of DNA was mixed in a 1.5 ml Eppendorf tube with 5 µl of T7 forward primer. The plasmids were then sent in Eurofins Genomics for sanger sequencing.

Restriction reaction	**V (µl)**	**Ligation reaction**	**V (µl)**
Restriction enzyme 1	0.5	10x T4 DNA ligase buffer	1
Restriction enzyme 2	0.5	T4 DNA ligase	0.5
10x FastDigest Green Buffer	1	+Plasmid backbone	
+1µg plasmid DNA		+insert	
+H_2O (total V=10µl)			

Restriction reactions were incubated for 1 hour at 37 °C, and the products were run on a 1 % TAE gel at 120 V for 30 minutes. Correct bands were cut from the gel and purified with NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The correct insert and backbone were then ligated together with total DNA amount of 50 ng. The individual amount of insert and plasmid DNA was determined by the ratio of their base pair amount. Ligation reaction was done at 4 °C for 24 hours.

3.10 Naïve reversion

Naïve reversions were done following the Naïve cult induction kit (Stemcell Technologies) protocol, and the induction mediums were also prepared with the medium components provided in the kit. NaB was used as HDACi during the protocol, and NaB concentration was optimized prior to the reversions by treating H9 PSCs with 0.5-1mM of NaB for 24 hours. A pilot reversion was done with H9 PSCs using different starting cell densities of 5 x 10^4, 1 x 10^5 and 2 x 10^5 cells. Later

reversions were done with densities of 5 x 10^4 and 2 x 10^5 cells using both H9 PSCs and the H9 DDdCas9VP192-LEUTX activator cell line (Table 8). Prior to the reversion all PSCs were seeded on iMEF feeder layer and cultured in the E8 medium supplemented with 10 µM ROCKi Y-27632 for 24 hours. The activator cell lines were treated with 0.5 µg/ml DOX (Sigma) and 2 µM TMP (Sigma) for 48 hours to induce LEUTX expression. Activator cell lines containing LEUTX gRNAs were reversed with the initial endogenous LEUTX expression and without as a control. Reversions were done in a hypoxic 37 °C, 5 % O_2, 5 % CO_2 growth conditions in X-vivo workstation. A single reversion of H9 cells was done in the atmospheric 20 % O_2 conditions.

Table 8. Naïve reversion sets

Naïve reversion	**Reverted cell lines**
Pilot reversion	H9 in 5 x 10^4, 1 x 10^5 and 2 x 10^5 cell densities
Reversion 1	H9, H9-DDdCas9-LEUTX +/- DOX/TMP in 5 x 10^4 density
Reversion 2	H9, H9-DDdCas9-LEUTX +/- DOX/TMP in 2 x 10^5 density
Reversion 3	H9 at 20 % O_2 in 2 x 10^5 density

3.5 cm culture plates were coated with 3 x 10^5 iMEFs during the reversion. During the first 3 days of the protocol the medium was supplemented with 0.5 mM NaB, acting as a HDACi. Depending on the colony size observed by microscope, the cells were passaged every 3-6 days starting from the day 10 of protocol. Cells were passaged as single cells with TrypLE, and the new medium was also supplemented with 10 µM ROCKi Y-27632 for 24 hours after the passaging. The reversion protocol took 6-9 weeks to complete, and at the end of the protocol the cells were transferred into t2iLGö medium after all of the background cells had disappeared. Reverted PSCs could be re-primed by culturing them for seven days on matrigel coated plates with E8 medium. Clones were frozen from all of the completed reversions.

3.11 Embryoid body assay

PSCs were passaged to ultra-low attachment 6-well culture plates and the medium was changed into a human embryonic stem cell medium hES without bFGF. Medium was also supplemented with 10 µM ROCKi Y-27632 for the first 24 hours. The PSCs were cultured as aggregates for two weeks. Every other day 2 ml of medium was added to culture plates, and the entire medium was changed every other day by centrifuging the EBs at 200 g for 2 minutes and resuspending them in

2 ml of hES without bFGF. After two weeks the EBs were plated into a 24 well-plate and cultured for 5-7 days in hES without bFGF. When the EBs had spread enough on the wells they were fixed for immunocytochemistry.

3.12 Alkaline phosphatase staining

The viability of NaB used in the reversions was tested with NaB-inducible human foreskin fibroblasts (HFF) used for iPSC reprogramming. HFF-retro-OK2nG-SMG cell line was cultured in hES medium for nine days with half of the culture plate wells induced with 0.25 mM NaB. After nine days the cells were alkaline phosphatase (AP) stained, and the number of formed iPSC colonies were compared (Supplementary Figure 2). The cells were first washed with PBS and then fixed with 4 % PFA in EDTA. 1 ml of AP solution was added into the wells, until background started to appear. AP-solution was then aspirated, and the cells were washed with PBS, and the final PBS was left to wells to prevent the cells from drying.

AP-solution	**V(µl)**
1 M Tris-HCl (pH 9.5)	100
5 M NaCl	20
1 M $MgCl_2$	50
NBT/BCIP (Sigma-Aldrich)	20
H_2O	810

4.Results

4.1 Selection of generated H9 activator cell line

H9 cells electroporated with plasmids containing the DDdCas9VP192 activator and rtTA were first evaluated by their GFP expression. Since DDdCas9 plasmid contained the EGFP coding sequence, the cells that had successfully integrated the DDdCas9 plasmid expressed GFP when treated with 0.5 µg/ml DOX (Figure 2). The GFP expression was heterogenous in the cell colonies and the intensity of the expression varied between cells.

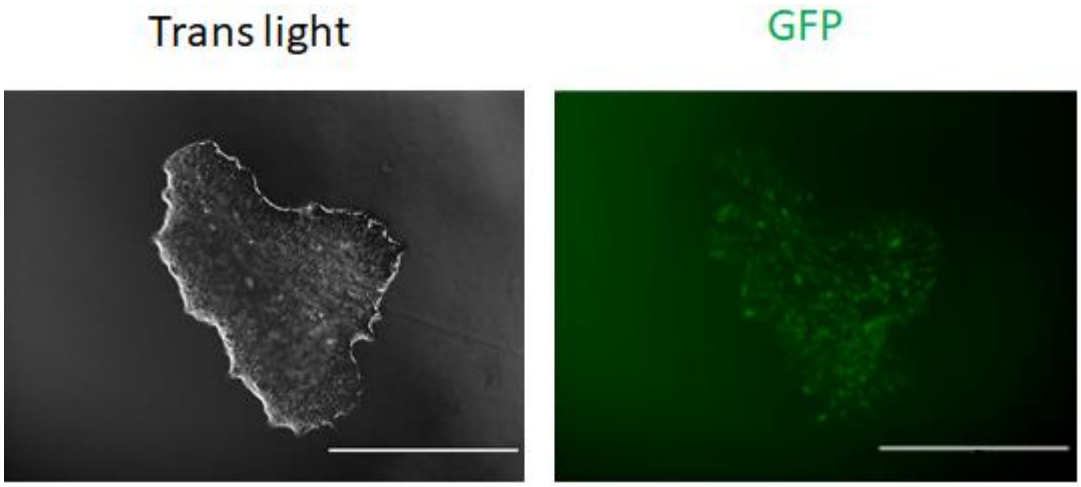


Figure 2. Electroporated H9 expressing GFP. Scale bars= 400 µm

Based on the results from the antibiotic concentration optimization done on H9 cells, the electroporated cells were selected with antibiotic treatment of 100 µg/ml G418 for 24 hours, followed by five-day treatment with 50 µg/ml G418. The remaining cells were treated with 0.5 µg/ml DOX in order to activate the activator plasmids TetON-promoter and zeocin resistance. The cells were then further selected with 1 µg/ml zeocin for 48 hours. The surviving cells, which had both plasmids successfully integrated, still expressed GFP and were used as an activator cell line for the naïve reversions.

4.2 Activator cell line test with endoderm factor guides

H9 activator cell lines ability to activate endogenous target expression was tested by transfected plasmid containing five guides for both FOXA2 and SOX17. After electroporation the cells were selected using 0.5 µg/ml puromycin antibiotic treatment for one day. After three-day treatment with 0.5 µg/ml DOX and 2 µM TMP the morphology of the cells started resembling endoderm cells. SOX17 was expressed in both RT-PCR and in immunostainings within the nucleus of the DOX/TMP treated samples, but not in the samples without DOX/TMP treatment (Figure 3).

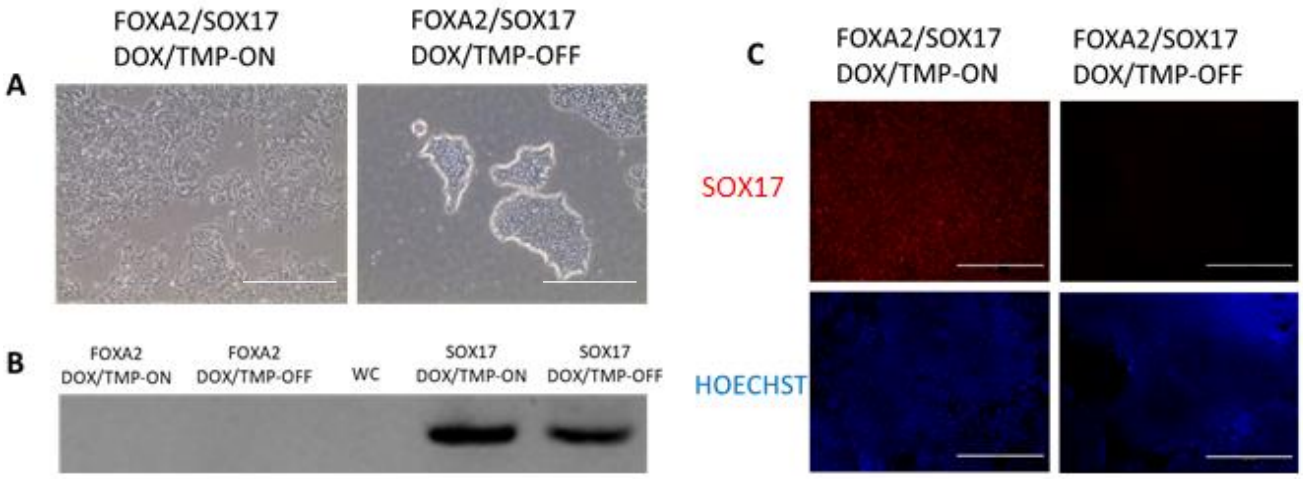


Figure 3. A) Morphology of H9 DDdCas9VP192-FOXA2/SOX17 cells with and without the three day DOX/TMP treatment. Scale bars= 400 µm

B) RT-PCR bands for FOXA2 and SOX17, WC= Water control

C) SOX17 immunostainings with and without DOX/TMP treatment. Scale bars= 400 µm

FOXA2 however was not expressed in any of the samples. The guide plasmids inserts were checked by EcoRI and NotI, as well as with SalI and SpeI digestions. The received bands did not match the expected lengths, indicating that there was something wrong with the guide plasmids assumed insert sequences.

4.3 Comparison of LEUTX gRNAs

After the four different LEUTX guide oligos were cloned into guide cassettes by guide PCR and then transfected into HEK293 cells, the cells were fixed and immunostained after three days with LEUTX antibody. In the immunostaining results the LEUTX expression of guide 5 stood out from the rest, so for the RT-qPCR assay HEK293 cells were also transfected with combinations of guide 5 and the other guides. Based on the RT-qPCR results the combination of LEUTX guides 4 and 5 was selected to be used as the guides for the H9 activator cell line (Figure 4).

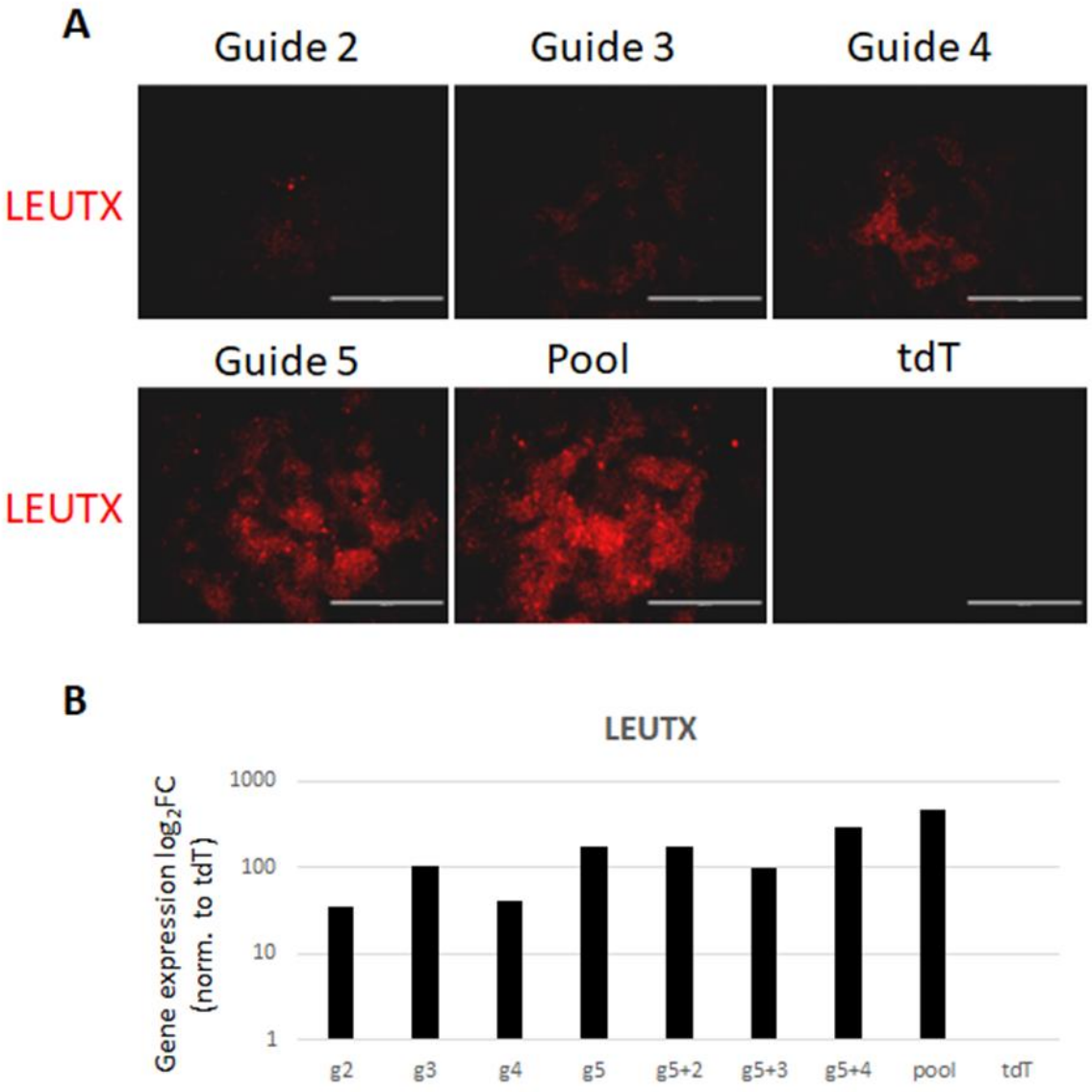


Figure 4. A) LEUTX guide HEK293 immunostainings with controls. Scale bars= 400 µm

B) LEUTX guide RT-qPCR. Expressions are normalized to the tdT. N=1

4.4 Verifying the assembled LEUTX gRNA plasmid

The LEUTX guide 4 and 5 cassettes assembled into GG-dest backbone by Golden Gate assembly were transformed into NEB5s and minipreps were made out from four picked colonies. Minipreps were then checked by EcoRI and NotI digestion and two of the digested bands looked to be the right size, so glycerol stocks were made from these two clones and samples were sent for sanger sequencing. One of the sequenced clones had the correct sequences for the both guides and also for the U6 promoter regions, so the guide insert was ligated into GG-dest-EBNA backbone, and later digested and ligated to the PB-GG-backbone in order to be electroporated into the activator cells.

After the ligation correct insert and plasmid backbone bands were verified with EcoRI and SpeI digestion (Figure 5).

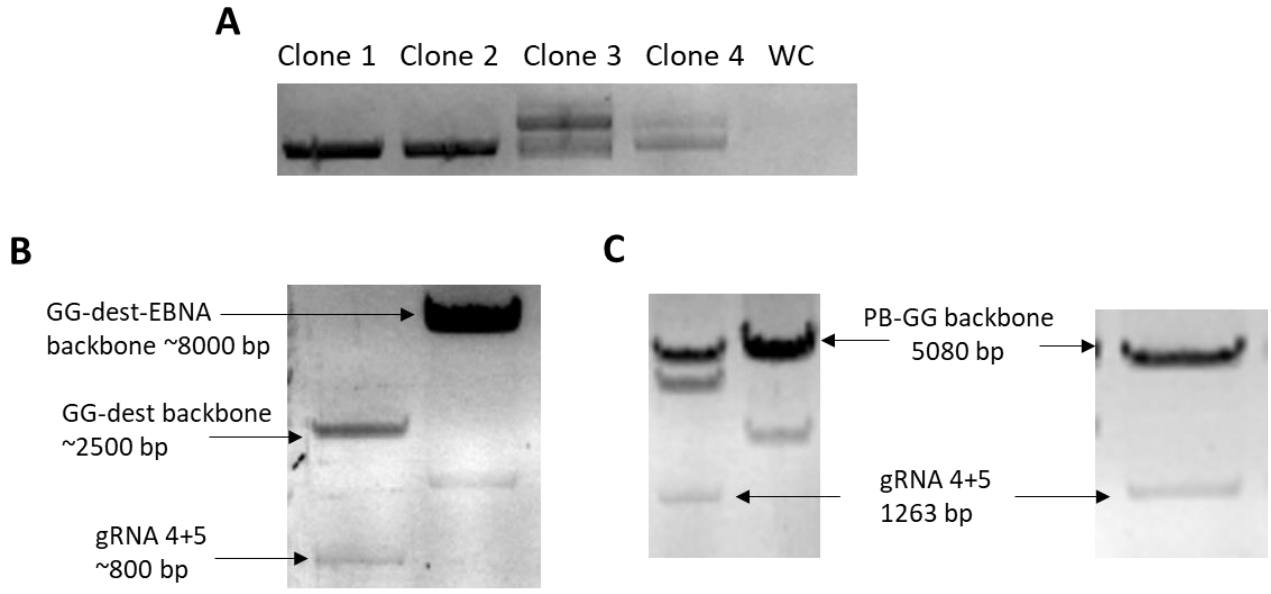


Figure 5. A) Digested minipreps from four picked guide plasmid clones. WC= Water control

B) gRNA insert and plasmid backbones after EcoRI and NotI digestion

C) Correct gRNA inserts and PB-GG-backbone before (left) and after (right) ligation

4.5 dCas9 and LEUTX expression in activator cells

After electroporating the finished PB-GG-LEUTXg4+5-PGK-Puro into H9 activator cells they were selected with 0.5 µg/ml puromycin for which the vector plasmid contained an antibiotic resistance. Selected activator cells were immunostained with LEUTX antibody after 1-3 days of 0.5 µg/ml DOX and 2 µM TMP treatment. Activator cells without DOX/TMP treatment were used as a control. LEUTX was expressed in nucleus of the H9 activator cells treated with DOX/TMP, but only faintly as background in the untreated cells. The expression was heterogeneous, with some individual cells expressing LEUTX stronger than others. There was no noticeable difference between the treated day 1-3 samples (Figure 6).

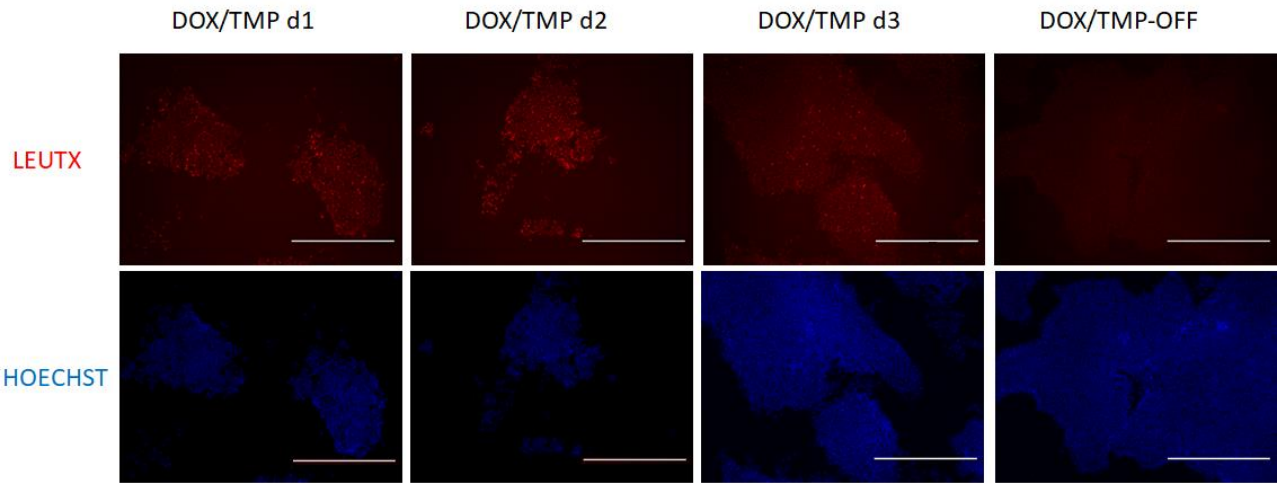


Figure 6. LEUTX expression in H9-DDdCas9VP192-LEUTX activator cells after 1-3 day DOX/TMP treatment. Scale bars= 400 µm

RT-qPCR assay was done with samples collected after 1-4 days of induced LEUTX activation by DOX/TMP treatment, with untreated cells and a water control as negative controls. Primers for both LEUTX and dCas9 were used in the RT-qPCR assay, and the assay was repeated with replicates. LEUTX expression peaked in the activator cells after a single day of DOX/TMP treatment and regressed on the following days while still being expressed. LEUTX expression was almost nonexistent in the untreated H9 activator cells. dCas9 expression was similar to the LEUTX expression, peaking at the day 1 and then regressing slightly. However, there was also some leaking dCas9 expression in the untreated activator cells, but the LEUTX expression remained low (Figure 7).

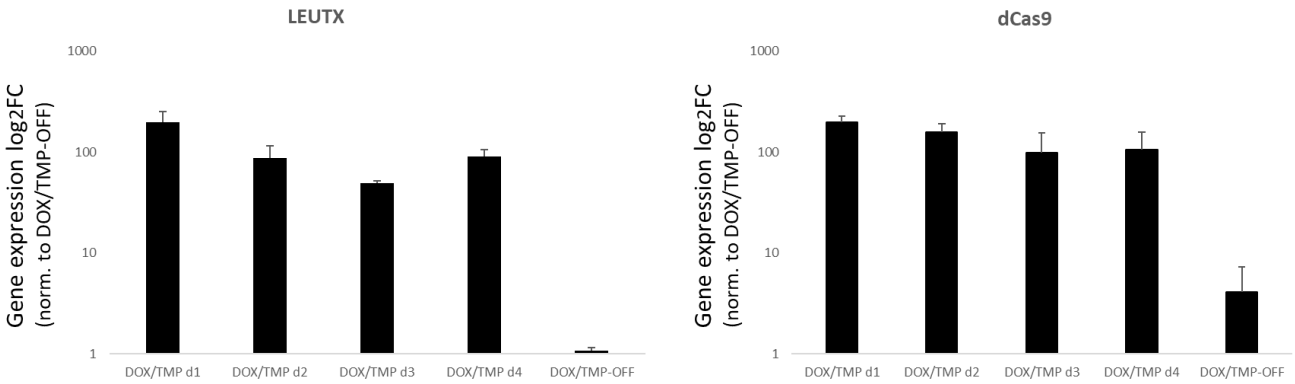


Figure 7. Relative LEUTX and dCas9 expressions in H9 activator cells. Both expressions are normalized on untreated DOX/TMP-OFF sample. Error bars represent standard deviation. N=2

4.6 Optimization of reversion protocol

The ideal concentration of NaB to be used at the start of the reversion protocol was tested on H9 cells with concentrations of 0.5-1 mM for 24 hours. With the concentrations of 0.75 mM and 1 mM a large amount of cell death was observed. With 0.5 mM the amount of cell death was much lower, so a NaB concentration of 0.5 mM was selected to be used at the reversions. NaB viability was also tested on NaB-inducible HFFs during fibroblast reprogramming. After 10 days the cells were AP-stained, and it was observed that the cells induced with NaB managed to form considerably more iPSC colonies compared to the non-induced ones.

First set of pilot reversions were done with three 3.5 cm culture plate wells of H9 cells in order to find an optimal starting cell density, and to practice the reversion protocol. The cells were seeded on iMEF feeder layer at different cell densities of 5 x 10^4, 1 x 10^5 and 2 x 10^5 cells. During the initial resetting phase H9 cells were induced with 0.5 mM NaB acting as HDACi on induction medium 1. The cells cultured on iMEF feeder layer in 5 % O_2 hypoxic growth conditions did not experience NaB caused cell death like during the optimization of NaB concentration.

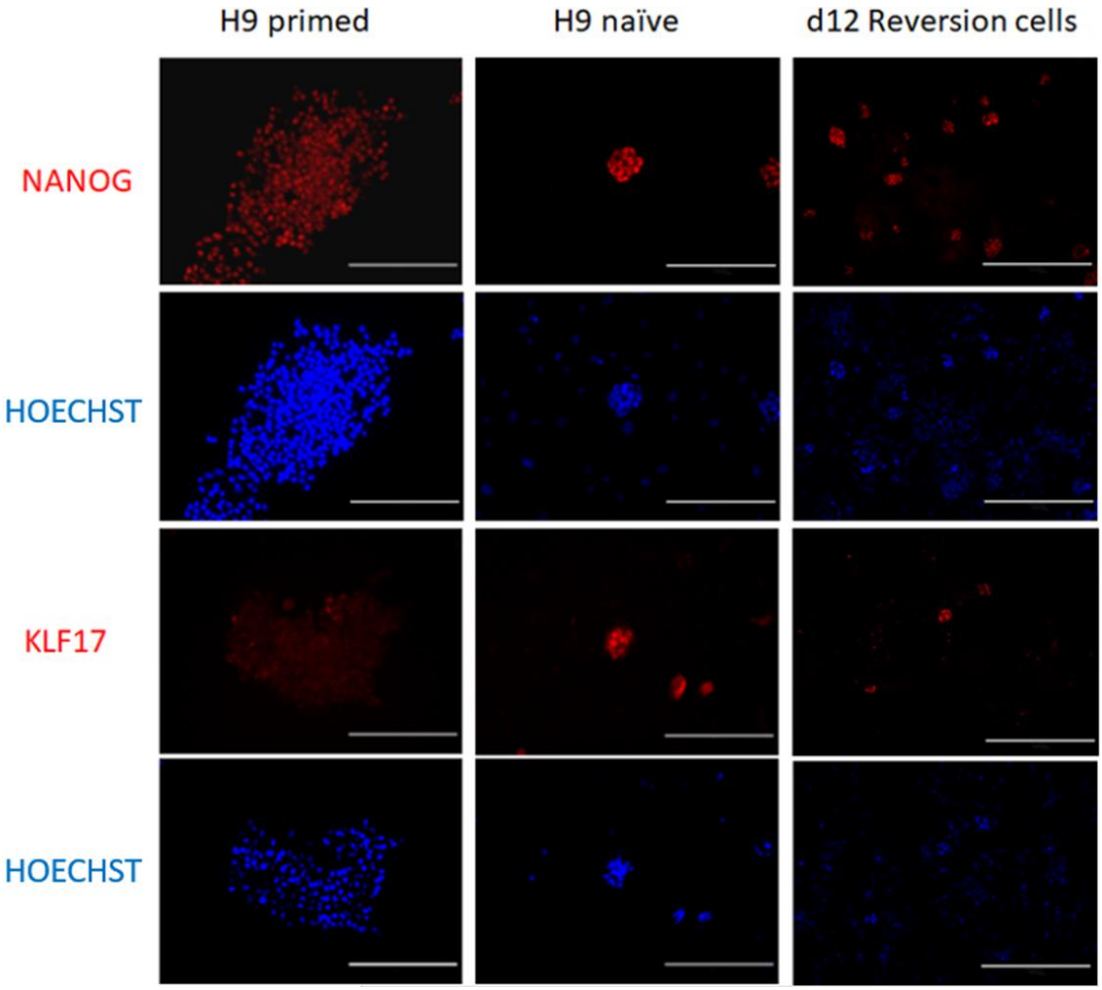


Figure 8. NANOG and KLF17 expressions in immunostained H9 primed, H9 naïve and day 12 reversion samples. Scale bars= 200 µm

All of the culture wells grew to full confluency before the first passaging at day 10, despite of the starting cell density. Even in the confluency the cells however managed to survive without any major cell death occurring. All wells started to successfully form naïve PSC clusters during induction medium 2, and samples for immunostainings were collected on day 12 after the first passage. The small cell clusters already expressed both NANOG and KLF17 naïve markers in nucleus at day 12. Both H9 primed and naïve PSCs were used as controls in the immunostainings. Both controls expressed NANOG in the nucleus, but KLF17 was only expressed in the nucleus of naïve H9 cells (Figure 8).

4.7 H9 and activator cell line reversions

After the optimization and pilot reversion set a total of three different reversion sets were done. The first set was done with one culture well of primed H9 cells, H9 activator cells with induced activation of endogenous LEUTX expression by 48 hour DOX/TMP treatment and H9 activator cells without the induced LEUTX expression. The second set was a replication of the first set with higher starting cell density. On the third reversion set a single well of primed H9 cells were reverted in atmospheric 20 % O_2 conditions.

In the first reversion set all of the cells expanded during induction medium 1, and after changing the medium to induction medium 2 naïve PSC clusters started to appear. The H9 activator cells with induced LEUTX expression prior to reversion formed almost four times more naïve clusters than H9 cells or the activator cells without induced LEUTX expression. Naïve clusters were manually counted from whole-well images prior to the second passage on day 16 (Figure 9).

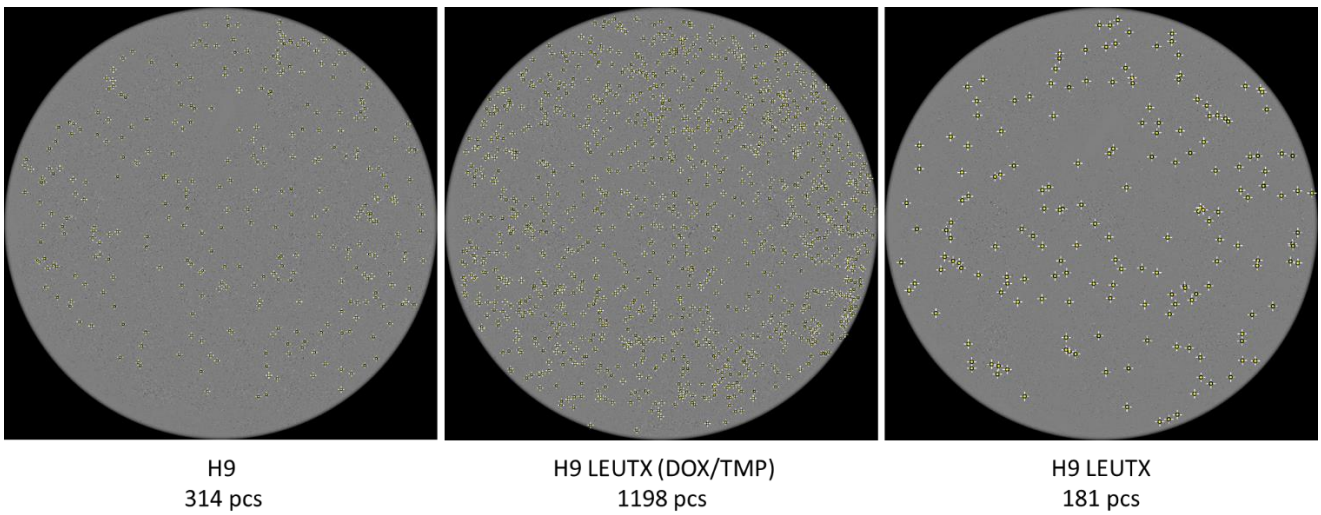


Figure 9. IncuCyte whole-well images of day 16 reversion, counting naïve clusters on H9 and H9-LEUTX activator cells treated with/without DOX/TMP prior to reversions

After changing the cells to induction medium 3, background cells started to disappear. However, the activator cells with induced LEUTX expression, which formed the most naïve colonies during induction medium 2, suffered from cell death in start of induction medium 3. This brought the amount of naïve colonies to the same level with H9 and non-induced activator cells. All cell lines however managed to complete the reversion protocol, and after changing the medium into t2iLGö samples were collected for RT-qPCR and immunostainings. In order to isolate enough RNA for cDNA synthesis, naïve samples had to be pooled from multiple culture wells.

At the start of the second reversion set H9 activator cells with induced activation of endogenous LEUTX expression prior to the reversion suffered from cell death after the first day of HDACi induction. However, like in the previous reversion these cells managed to form naïve colonies more efficiently during the induction medium 2 and caught up with the other cells in the amount of formed naïve colonies. Otherwise there were no major differences compared to the first reversion set, so after the reversion was complete samples were collected for RT-qPCR and immunocytochemistry. In the final reversion set the H9 cells reverted in atmospheric 20 % O_2 conditions started experiencing cell death on induction medium 1. This was similar to the cell death observed during optimization of NaB concentration. During induction medium 2 the cells formed less naïve colonies when compared to the earlier reversions, and the non-reset background cells started to differentiate easily (Figure 10).

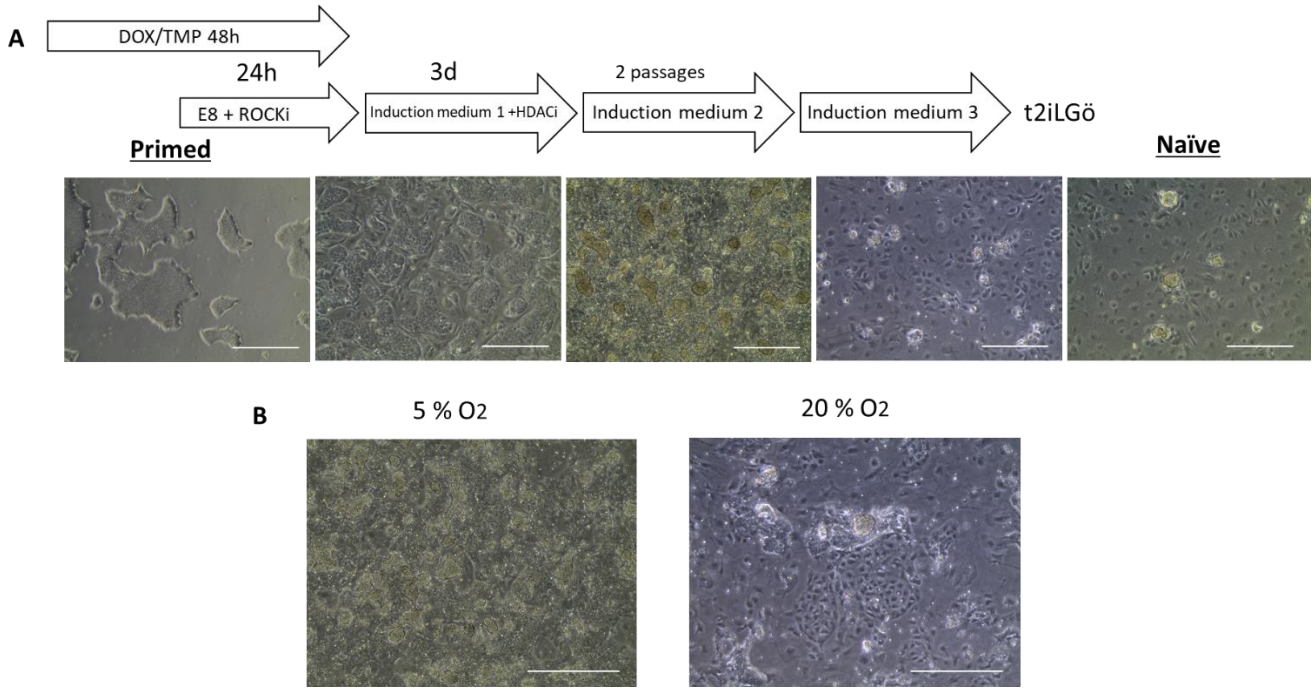


Figure 10. A) Images of H9 cells during the reversion protocol. Scale bars= 400 µm

B) H9 cells in 5 % and 20 % O_2 during day 8 of the reversion. Scale bars= 400 µm

The number of differentiated cells decreased during induction medium 3, and the amount of naïve colonies was expanded. The reversion was completed successfully, but the cells were prone to easily re-prime and differentiate while being expanded on the t2iLGö medium. Samples were collected for the RT-qPCR and immunocytochemistry.

All of the completed reversions and the deviations observed in reversions of induced activator cells and H9 cells reverted in 20 % O_2 are summarized in table 9. Completed reversion cells were characterized based on their morphology and naïve marker expressions in immunocytochemistry and RT-qPCR assays.

Table 9. Summary of the completed naïve reversions

Cell line	**DOX/TMP induction**	**Starting cell density**	**O_2 c (%)**	**Notes**
H9	No	5 x 10^4	5	
H9	No	2 x 10^5	5	
H9	No	2 x 10^5	20	Reduced naïve colony formation, cell death observed on induction medium 1, increased amount of cell differentiation
H9-DDdCas9-LEUTX	Yes	5 x 10^4	5	Increased amount of naïve colony formation on induction medium 2
H9-DDdCas9-LEUTX	Yes	2 x 10^5	5	Cell death observed on day 1 of the protocol, increased amount of naïve colony formation on induction medium 2
H9-DDdCas9-LEUTX	No	5 x 10^4	5	
H9-DDdCas9-LEUTX	No	2 x 10^5	5	

4.8 Characterization of reverted PSCs

Reverted PSCs had a clear difference in cell morphology, with primed PSCs prior to the reversion growing as a flat monolayer, while the naïve control PSCs grew in small domed clusters. All of the reverted PSCs had the same domed morphology similar to the control naïve H9 PSCs (Figure 11).

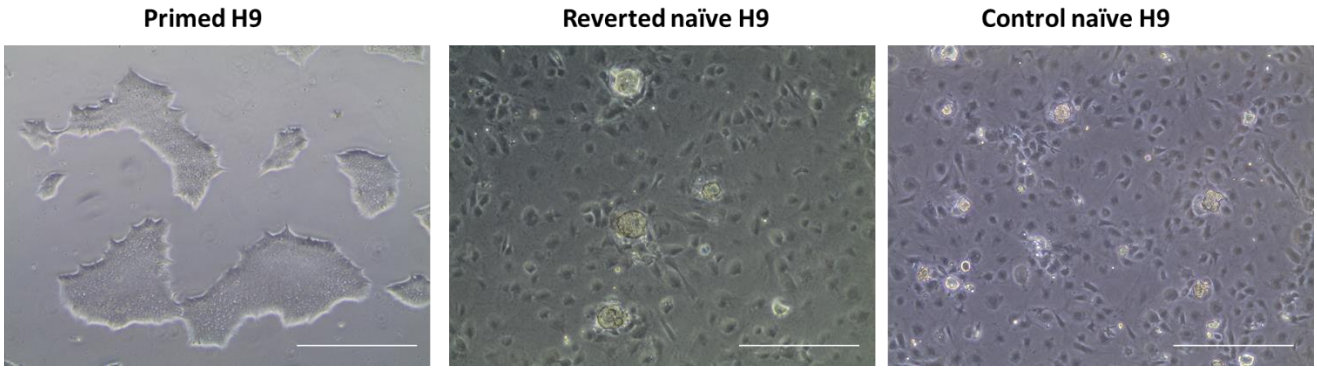


Figure 11. Primed and naïve H9 morphologies. Scale bars= 400 µm

Immunostainings were made from all of the collected reversion samples with NANOG and KLF17 antibodies. Expression of both markers was present in the nucleus on all of the reverted PSCs similar to the expression observed in control naïve H9 PSCs (Figure 12 A). Reverted PSCs, as well as the naïve and primed H9 controls, were also immunostained with H3K27me3 methylation marker. However, the expression was only observed with high magnification on a couple of individual primed control PSC nuclei as small dots (Figure 12 B). qPCR primers were tested on primed and naïve H9 control PSCs prior to the RT-qPCR assay. Original KLF17 primers had a mistake in the reverse primer sequence so a corrected primer, as well as a pair of new KLF17 primers, were designed with primer BLAST. DNMT3L primers also showed primer dimers in the primed H9 controls, so a new pair of primers were designed and order at the same time with the KLF17 primers. The new designed primers were then used for the RT-qPCR assays.

In the RT-qPCR assay primers for naïve markers NANOG, KLF17, DNMT3L and TFCP2L1 were used. In addition to the collected reversion samples, primed and naïve H9 samples were used as a control. Fold change in the expression of NANOG was lower between the naïve and primed PSCs when compared to the rest of markers. NANOG expression was still higher in all of the reverted PSCs and in the naïve control when compared to the primed control, except for the H9 PSCs reverted in 20 % O_2. For the rest of the marker's expression was always much higher in the

reverted PSCs and in naïve control, especially with DNMT3L expression which was over 1000 times the fold change in naïve PSCs when compared to the primed PSCs (Figure 13).

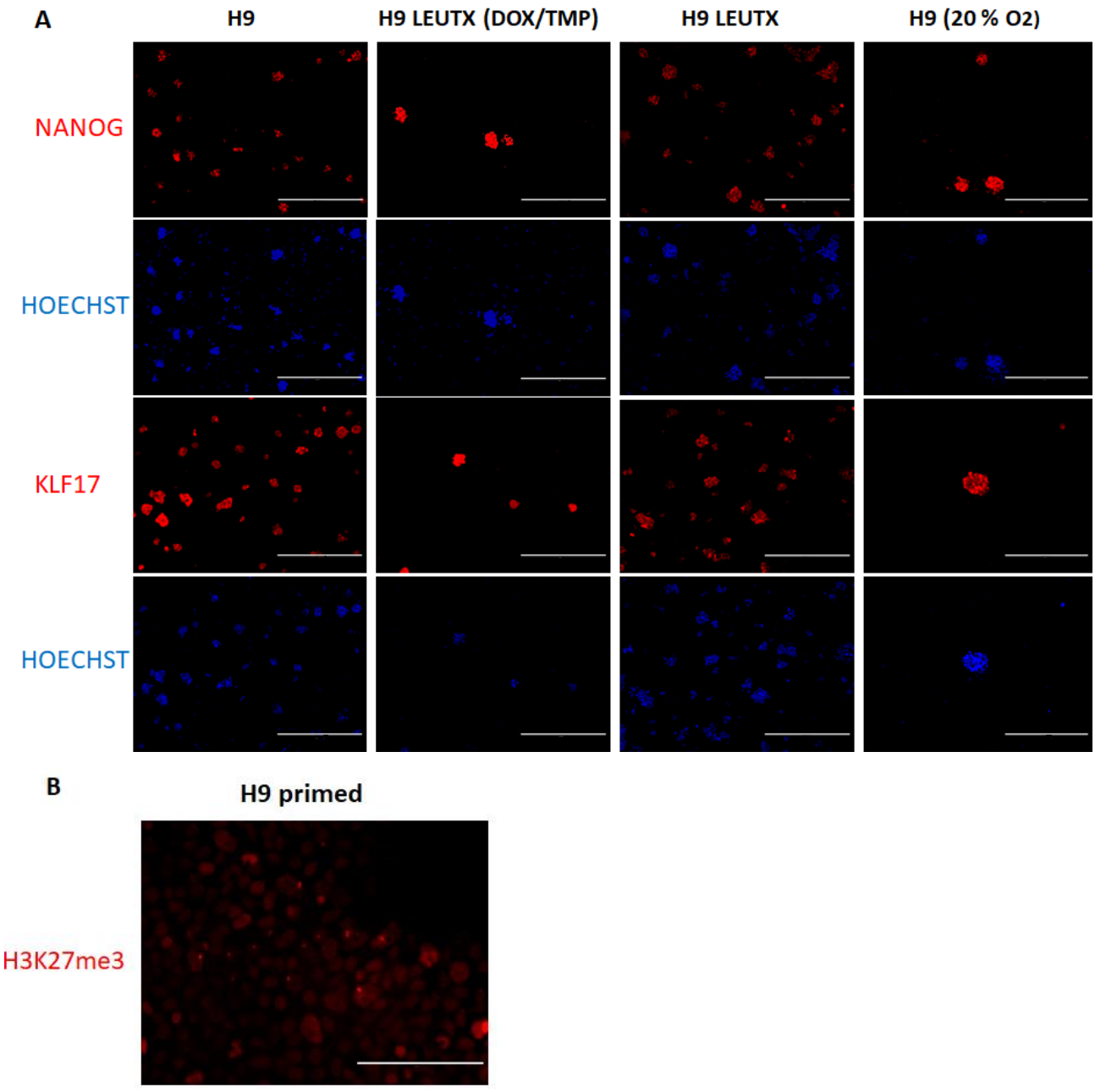


Figure 12. A) NANOG and KLF17 immunostainings of reverted PSCs. Scale bars= 400 µm **B)** H3K27me3 marker expression in primed PSCs. Scale bar= 100 µm

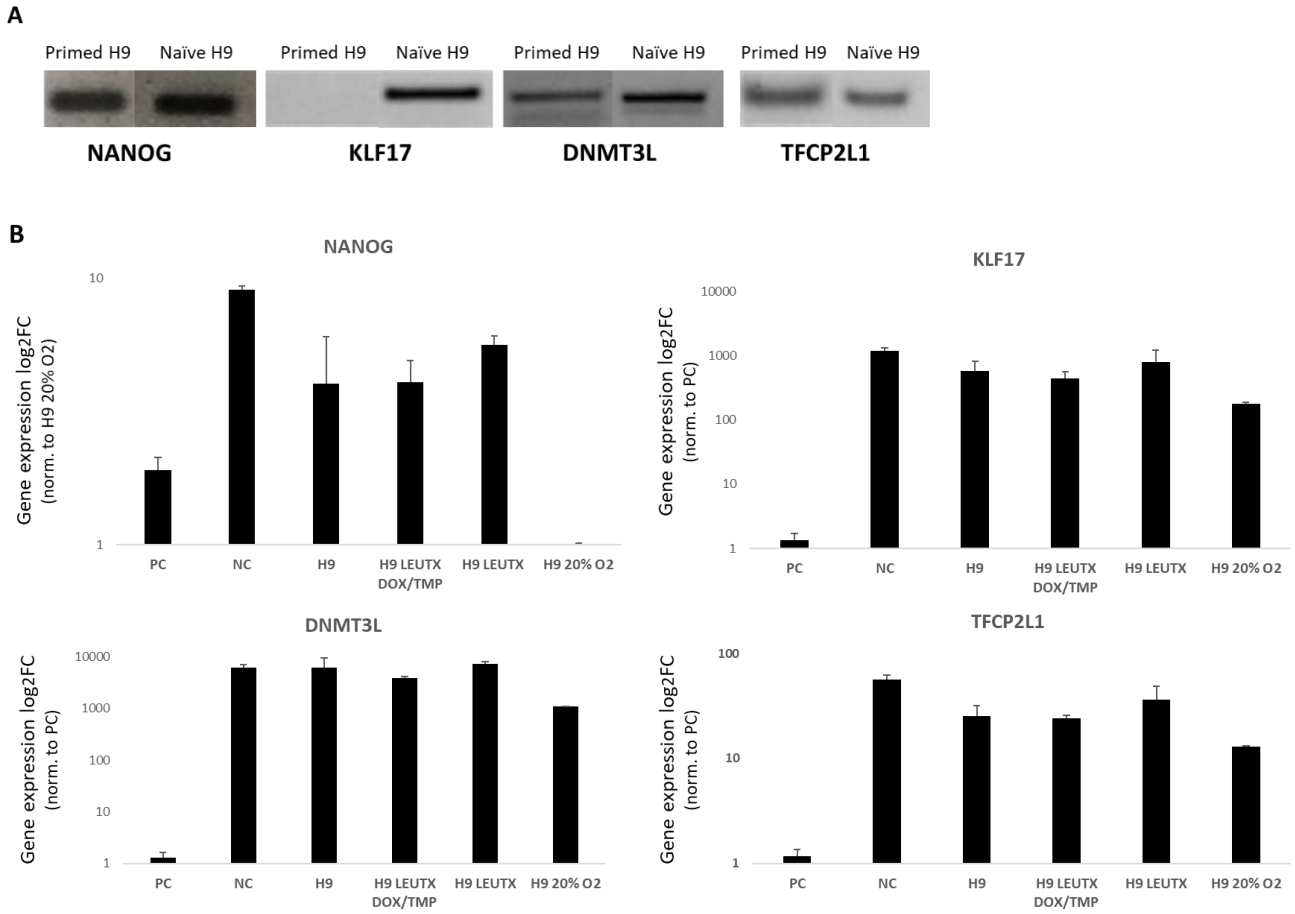


Figure 13. A) RT-PCR primer bands for naïve markers

B) Relative gene expressions in reversed H9 and H9-LEUTX cells with and without DOX/TMP treatment. PC=Primed H9 control, NC= Naïve H9 control. NANOG is normalized on H9 20% O_2, other markers on PC. Error bars represent standard deviation. N=2

Based on the results of reverted PSCs morphology, and their naïve marker expressions in immunocytochemistry and RT-qPCR assay, the reverted PSCs were characterized as naïve PSCs.

4.9 Embryoid body formation

EBs were made from both primed H9 and H9 activator cells. Reverted naïve H9 PSCs failed to form EBs and were first required to be re-primed back into the primed state. H9, H9 activator cells and re-primed H9 PSCs successfully formed EBs. H9 activator EBs were immunostained with endodermal marker SOX17, ectodermal marker β-tubulin III and with mesodermal marker α-Smooth muscle actin demonstrating their capacity to form into all three germs layer derivatives. (Figure 14).

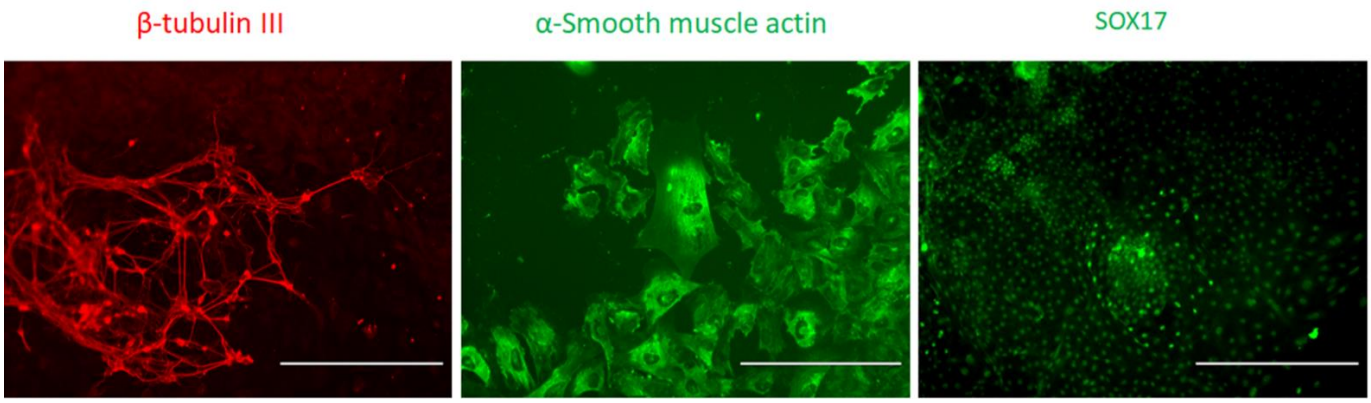


Figure 14. Immunostainings for EBs formed from H9 activator cells. Scale bars= 400 µm (β-tubulin III, SOX17), 200 µm (α-Smooth muscle actin)

5. Discussion

5.1 Generation of functional H9 activator cell line

The inducible activator cell line was generated successfully by integrating the DDdCas9VP192 activator construct described in Balboa et al. 2015 via SB transposition to H9 PSCs. The cells maintained their primed morphology, and when cultured they were capable of self-renewal and proliferation for multiple passages without starting to differentiate. When compared to the H9 cells however the activator cells were more fragile and prone to differentiation and cell death especially when the cell confluency became too high on the culture plates. The plasmid containing gRNAs targeting FOXA2/SOX17 used to test the activator cell line's functionality did not have the expected inserts when their plasmids were digested, which could explain why FOXA2 expression failed to activate. Endogenous SOX17 expression was however activated successfully, indicating that the error in the plasmid insert was more likely amongst the five FOXA2 guides inserted into the plasmid. Endogenous SOX17 activation was still enough for the cells to start differentiating towards endodermal lineage, which was observed by the change in the cell's morphology as expected (Shimoda et al. 2007). Furthermore, endogenous SOX17 was not expressed in the untreated cells, indicating that the TetON and DHFR DD inducible systems were functioning as expected and preventing the unwanted dCas9 leakage and target activation even when the guide plasmid was present in the cells (Balboa et al. 2015; Weltner et al. 2018). The pluripotency of activator cell line was verified by their spontaneous differentiation into derivatives of all three germ line lineages in the EB assay.

5.2 Endogenous LEUTX expression can be activated by gRNA targeting

FuGENE HD chemical transfection of HEK293 cells was used for the LEUTX gRNAs testing purposes since it is an easy and cheap way to transfect smaller vectors into the fast proliferating HEK293 cells, which are efficient for transfections and protein production (Thomas and Smart 2005). Electroporation on the other hand is more efficient way of transfecting ESCs (Tabar et al. 2015) so therefore it was used for all activator cell line transfections. There was also no need to integrate the activator nor the guides into the genome for the testing, so an episomal pCXLE-activator plasmid and purified guide PCR products were used. However, it should be noted that in the RT-qPCR assay the pool sample containing all four guides had an even higher LEUTX expression than the two-guide combination. We still decided to use the two-guide combination for the reversion experiments because the LEUTX expression was almost at the same level as with the four-guide pool sample, and during the time four-guide combinations were already being tested in

other dCas9 activator cell line experiments performed within our research group. Therefore, we wanted to see if the two-guide combination was sufficient to activate the target gene expression in the activator cells.

After the guide plasmid was Golden Gate assembled and eventually transfected into the activator cells, the gRNAs were able to successfully target and activate the endogenous expression of LEUTX. The activator cells could also still be maintained and expanded after the guides were electroporated into the cells. LEUTX was not expressed without DOX/TMP induction, but there was some dCas9 expression in the cells even without DOX/TMP. Apparently, the small amount of leaky dCas9 is not enough to activate the endogenous expression of the target gene as it is most likely highly unstable without the presence of TMP. Higher dCas9 expression levels also resulted in higher LEUTX expression, which makes sense since the dCas9VP192 targets and activates the *LEUTX* transcription with the help of electroporated guides. We were expecting that the dCas9 and LEUTX expression levels would gradually increase during the four-day treatment with DOX and TMP, but instead the expression levels peaked already after one day of treatment. The expression levels between the four days were still on similar levels and the immunostainings of samples from different days of DOX/TMP induction looked also similar. For activator cells expressing endogenous LEUTX in the naïve reversions, we chose to use 48 hour DOX/TMP induction prior to the start of the reversion, so that the activator cells could maintain the LEUTX expression for a while. Since the expression levels rose quickly after one day DOX/TMP treatment, they could also decrease rather quickly without longer DOX/TMP induction prior to the reversions. This could be measured by collecting samples after the DOX/TMP treatment has been stopped for further optimization in the future experiments.

5.3 Naïve colonies formed during pilot reversion

The 0.5 mM concentration of NaB was selected for the reversions because of the amount of cell death observed in higher concentrations. However, when the pilot reversion was done with H9 cells starting in three different cell densities there was no noticeable cell death present. On the contrary the cells survived and continued to expand even in high cell confluency regardless of the starting cell density. Since the reversions are done on iMEF feeder layer, they can secrete growth factors into the medium and improve the proliferation and survivability of the cells even in full confluency (Llames et al. 2015). In the future experiments higher NaB concentrations of 0.75 and 1 mM could be also tested to see if this would improve the initial resetting of the PSCs and have a positive effect on the reset PSCs ability to generate naïve colonies during induction medium 2. In order to make

sure that there was nothing wrong with the NaB itself it was tested on the NaB-inducible HFFs. Since the NaB treated HFFs ended up forming more iPSC colonies than the untreated ones, we could conclude that NaB was working as expected. After the initial resetting with HDACi, the medium was changed into induction medium 2 and the cells started to form naïve colonies just as described by the Naïve cult induction kit's reversion protocol. The formed naïve colony clusters had a domed morphology, as described in previous findings (Guo et al. 2017), while the non-reset cells remained as background. After immunostaining the cells on day 12 the colonies were already expressing naïve markers NANOG and KLF17. This was to be expected since according to the previous findings NANOG expression is increased by over 2-fold within the first nine days, and the KLF17 expression should become noticeable after day seven (Guo et al.2017).

5.4 Induced activator cells improve the formation of naïve-like colonies

During the reversions H9 PSCs and the H9 activator cells without the DOX/TMP induction did not have noticeable differences, and they both managed to successfully go through the entire reversion protocol and form stable naïve PSCs. The reversions were successful with both starting cell densities. Since the reversions done with 2×10^5 starting cell density did not suffer from cell death even while in full cell confluency, they had more overall cells on culture plates when the medium was changed into induction medium 2. Thus, when the naïve colonies started to appear they formed more naïve colonies when compared to the 5×10^4 starting density. The H9 activator cells with endogenous LEUTX expression starting in 2×10^5 cell density suffered from cell death after the first day of HDACi treatment. Since this was not observed in the other reversions, including the H9 activator cells with endogenous LEUTX expression starting in 5×10^4 cell density, the cell death was most likely caused by a human error during the seeding of cells on iMEF layer, or in the concentration of DOX, TMP or HDACi. After seeding the cells on iMEF feeder layer 24 hours before the initial HDACi resetting phase, the cells are also supplemented with 10 µM ROCKi Y-27632 in order to improve survivability during single cell passaging. A mistake in ROCKi concentration could also explain the observed cell death. The reversion with induced activator cells should be repeated in order to get reliable replicated results.

All of the H9 activator cells with endogenous LEUTX expression however managed to also complete the reversion protocol and formed stable naïve PSCs. The activator cells with endogenous LEUTX expression did form almost four times more naïve colonies during induction medium 2, but most of these colonies were lost after changing the medium into induction medium

3. Induction medium 3 is not supplemented with XAV939, and it's during this point when the non-rest background cells also disappear. Since LEUTX activates multiple genes during the EGA in pre-implantation embryos, it is possible that these genes could be involved in generation of both the ICM and trophoblast lineages. Also, in addition to the primed and naïve states of pluripotency there exists intermediate states that are still less well-known and characterized (Yang et al.2019). It is possible, that the transcriptome of at least some of the morphologically naïve-looking clusters formed by the LEUTX expressing activator cells during induction medium 2 varied from the naïve clusters formed in other reversions and resembled more of the less well-known intermediate states of pluripotency or the trophoblast lineage. Thus, these naïve looking PSCs were not able to survive after the transition to induction medium 3 containing small-molecule inhibitors for naïve PSC expansion, but still some of the naïve clusters formed by LEUTX expressing activator cells managed to survive and complete the reversion process. The reverted naïve PSCs were unable to form EBs, as previously reported (Guo et al. 2017). Naïve PSCs are not primed for differentiation, so after re-priming the naïve PSCs back to the primed state they were able to successfully form EBs in which the cells spontaneously differentiated into all three germ lineages. The re-priming takes considerably shorter period of time (7-10 days) compared to the naïve reversion (6-9 weeks) and requires minimal effort by just changing the culture medium. This is because the naïve PSCs represent an earlier stage in embryo development and are naturally being driven towards the primed state of pluripotency.

5.5 Oxygen concentration affects the reversion

The naïve reversions were done in 5 % O_2 concentration. Even though this was referred as hypoxia, the true hypoxia for ESCs is even lower around 2 % O_2 (Forristal et al. 2010). The reversion protocol instructed the use of 5 % O_2 hypoxic growth conditions which has also been used in the previous reversion experiments (Guo et al. 2017). To observe the effect of hypoxia on the naïve reversion, one reversion set was done in atmospheric 20 % O_2. These cells suffered from cell death during the HDACi resetting phase, formed less naïve colonies and differentiated more compared to the other reversions. The PSCs induced with multiple small-molecule inhibitors during the reversion protocol are already under cellular stress, so the added effect of high oxygen concentration clearly has a negative impact on the cells during the reversion. Since the high oxygen concentration is known to reduce proliferation and the expression of core pluripotency factors

(Forristal et al. 2010) it makes sense that the reversion efficiency is reduced, and the cells tend to easily differentiate even when expanded in t2iLGö medium after the reversion.

5.6 Naïve marker expressions

The differences in NANOG expression between primed and naïve samples were lower than with other markers, which was expected since NANOG is a pluripotency marker expressed in both PSC states (Nichols and Smith 2009; Chambers et al. 2003). The NANOG expression was still higher in naïve samples, with the exception of H9 PSCs reverted at 20 % O_2. As previously mentioned, high oxygen concentration is known to reduce the expression of core pluripotency factors, including NANOG (Forristal et al. 2010), which is supported by these results. In immunostainings NANOG was expressed in the nuclei of all of the samples just as expected. KLF17 was only expressed in the nuclei of naïve PSCs, and in RT-qPCR expressions fold change was over 100 times higher in all naïve samples compared to the primed PSCs. Out of the naïve samples the expression was lowest in the 20 % O_2 H9 reverted PSCs, but the expression level was still closer to the other naïve samples than the primed PSCs. The results are logical when considering that KLF17 is exclusively observed in primate ICM (Blakeley et al. 2015; Boroviak et al. 2015). The H3K27me3 marker was only observed in couple of primed H9 cells. The EVOS FL cell imaging systems magnification was not sufficient to reliably detect the H3K27me3 expression, so the characterizations were done only based on the cell morphology and the other naïve markers expressions. The detection of H3K27me3 could be improved by the use of more powerful confocal microscopes. The TFCP2L1 and DNMT3L expressions fold changes were much higher in naïve samples when compared to primed samples, as expected. This is supported by previous findings (Ye et al. 2014; Neri et al. 2013) where both markers are highly expressed in naïve PSCs. As with KLF17, the expression of TFCP2L1 and DNMT3L was lower in H9 PSCs reverted at 20 % O_2 when compared to the other naïve samples, but still higher than in primed PSCs. As previously mentioned high oxygen concentration can reduce expression of pluripotency factors, so it's unsurprising that the expressions of other naïve markers were also lower in H9 PSCs reverted at 20 % O_2 when compared to the other naïve samples.

5.7 The future prospects

Based on the morphology of the reverted PSCs and the marker expression levels, the reverted PSCs were characterized as naïve PSCs. The reverted activator cells that expressed LEUTX did not show noticeable differences in the marker expressions compared to the other reverted naïve PSCs, but the increased number of naïve-looking colonies during the induction medium 2 raises questions. Further experiments are required to better study the transcriptome of these forming cell clusters and their possible connection to the intermediate states or the trophoblast. Since endogenous LEUTX expression did improve the formation of these clusters during reversion, it would also be also interesting to see if LEUTX expression could potentially replace the requirement of the initial HDACi resetting phase at the start of the reversions. If HDACi enables the resetting towards the naïve state by making the chromatin more open and thus reducing the silencing of naïve factors (Guo et al. 2017), perhaps the EGA involved genes activated by LEUTX could help in unsilencing these factors without the additional need for HDACi. Collecting samples for immunostainings and RT-qPCR during the reversion is challenging due to the small number of cells in the forming naïve clusters, so a naïve cell-surface marker could help in characterizing the cells during the reversion. Recently a sushi containing domain 2 (SUSD2) was shown to co-localize with KLF17 in pre-implantation cells (Bredenkamp et al. 2019). As a cell-surface marker SUSD2 expression could be observed during the reversions without the need to fix the cells. By using the IncuCyte live-cell analysis system this would allow us to follow the change in naïve markers expression at the reversion cells on a daily basis. This would also help in characterizing the naïve-looking colonies formed by induced LEUTX expressing cells and enable the comparison of these clusters' marker expression to other naïve colonies in future experiments. The naïve PSCs reverted from the activator cell line still contain the DDdCas9 in their genome, which mean they can now be utilized to study the effects of target genes in human naïve PSCs. It would be interesting to see what effects the endogenous expression of LEUTX would have after reversion in finished naïve PSCs. Since LEUTX is mainly expressed during 4 and 8 cell stages (Jouhilahti et al. 2016) it could drive the naïve PSCs transcriptome towards even earlier cell stages.

The naïve reversion protocol was set up and optimized successfully and can now be used as a reliable way of obtaining human naïve PSCs for further experiments studying and modelling the earlier cell stages in embryo development. Furthermore, the conditionally stabilized DDdCas9 activator cell line worked as intended and can be utilized for studying the effects of other targeted genes during the reversion or on the reverted naïve PSCs.

6. Acknowledgments

First of all, thank you Timo Otonkoski for recruiting and giving me the change to work on this fascinating project as a part of your research group. It was truly an honor to get to work alongside within your prestigious group consisting of absolutely wonderful group members in a supportive atmosphere. And a huge thank you to my primary supervisor Ras Trokovic who gave me guidance throughout the entire project. Your positive attitude and supportiveness always ensured a nice working environment.

Additionally, thank you to my co-supervisors Eeva-Mari Jouhilahti and Jere Weltner, who both taught me various assays and answered all of my infinite questions. Your passion for the field is truly inspiring. Thank you, Eeva-Mari, especially for your help on generating the activator cell line and all things considering LEUTX. And thank you Jere for your help in CRISPR associated assays and with the plasmid constructions and cloning.

I would also like to thank Anni, Heli, Agnes and Jarkko for preparing the necessary materials and keeping the whole laboratory running. Thank you Anni for guiding me during the first weeks and helping me to get accustomed to the laboratory. Also, thank you Hazem for helping me with the X-vivo and IncuCyte and thank you Hossam for showing me the qPCR assay, even when no one else was around in the lab. And a special thanks to you Andrew, for your continuous support and friendship both inside and outside of the laboratory.

Finally, I am grateful to my family and friends for their continuous support during my master's thesis.

7. References

Amps, K., Andrews, P.W., Anyfantis, G., Armstrong, L., Avery, S., Baharvand, H., Munoz, M.B., Beil, S., Benvenisty, N., Baker, J., et al. (2012). Screening a large, ethnically diverse population of human embryonic stem cells identifies a chromosome 20 minimal amplicon that confers a growth advantage. Nat. Biotechnol. *29*, 1132–1144.

Balboa, D., Weltner, J., Eurola, S., Trokovic, R., Wartiovaara, K., and Otonkoski, T. (2015). Conditionally Stabilized dCas9 Activator for Controlling Gene Expression in Human Cell Reprogramming and Differentiation. Stem Cell Reports *5*, 448–459.

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature *461*, 1292–1295.

Barakat, T.S., Ghazvini, M., De Hoon, B., Li, T., Eussen, B., Douben, H., Van Der Linden, R., Van Der Stap, N., Boter, M., Laven, J.S., et al. (2015). Stable X chromosome reactivation in female human induced pluripotent stem cells. Stem Cell Reports *4*, 199–208.

ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nat. Cell Biol. *13*, 1070–1075.

Blakeley, P., Fogarty, N.M.E., Valle, I., Wamaitha, S.E., Hu, T.X., Elder, K., Snell, P., Christie, L., Robson, P., Niakan, K.K., et al. (2015). Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. Development *142*, 3151–3165.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat. Cell Biol. *16*, 516–528.

Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A., and Bertone, P. (2015). Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. Dev. Cell *35*, 366–382.

Bredenkamp, N., Stirparo, G.G., Nichols, J., Smith, A., and Guo, G. (2019). The Cell-Surface Marker Sushi Containing Domain 2 Facilitates Establishment of Human Naive Pluripotent Stem Cells. Stem Cell Reports *12*, 1212–1222.

Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes,

S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448*, 191–195.

Burdon, T., Chambers, I., Stracey, C., Niwa, H., and Smith, A. (1999). Signaling Mechanisms Regulating Self-Renewal and Differentiation of Pluripotent Embryonic Stem Cells. Cells Tissues Organs *165*, 131–143.

Bürglin, T.R. (2011). Homeodomain Subtypes and Functional Diversity. Subcell. Biochem. *52*, 95–122.

Cary, L.C., Goebel, M., Corsaro, B.G., Wang, H.G., Rosen, E., and Fraser, M.J. (1989). Transposon mutagenesis of baculoviruses: analysis of Trichoplusia ni transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. Virology *172*, 156–169.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional Expression Cloning of Nanog, a Pluripotency Sustaining Factor in Embryonic Stem Cells Ian. Cell Press Funct. *113*, 643–655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Chen, Y.T., Dejosez, M., Zwaka, T.P., and Behringer, R.R. (2008). H1 and H9 human embryonic stem cell lines are heterozygous for the ABO locus. Stem Cells Dev. *17*, 853–855.

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res. *23*, 1163–1171.

Davidson, K.C., Mason, E.A., and Pera, M.F. (2015). The pluripotent state in mouse and human. Development *142*, 3090–3099.

Dietrich, J.-E., and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. Development *134*, 4219–4231.

Dutta, D., Ray, S., Home, P., Larson, M., Wolfe, M.W., and Paul, S. (2011). Self-renewal versus lineage commitment of embryonic stem cells: Protein kinase C signaling shifts the balance. Stem Cells *29*, 618–628.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from

mouse embryos. Nature *292*, 154–156.

Forristal, C.E., Wright, K.L., Hanley, N.A., Oreffo, R.O.C., and Houghton, F.D. (2010). Hypoxia inducible factors regulate pluripotency and proliferation in human embryonic stem cells cultured at reduced oxygen tensions. Reproduction *139*, 85–97.

Fraser, M.J., Smith, G.E., and Summers, M.D. (1983). Acquisition of Host Cell DNA Sequences by Baculoviruses: Relationship Between Host DNA Insertions and FP Mutants of Autographa californica and Galleria mellonella Nuclear Polyhedrosis Viruses. J. Virol. *47*, 287–300.

Galliot, B., de Vargas, C., and Miller, D. (1999). Evolution of homeobox genes: Q50 paired-like genes founded the paired class. Dev. Genes Evol. *209*, 186–197.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. U. S. A. *89*, 5547–5551.

Gumireddy, K., Li, A., Gimotty, P.A., Klein-Szanto, A.J., Showe, L.C., Katsaros, D., Coukos, G., Zhang, L., and Huang, Q. (2009). KLF17 is a negative regulator of epithelial-mesenchymal transition and metastasis in breast cancer. Nat. Cell Biol. *11*, 1297–1304.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development *136*, 1063–1069.

Guo, G., Von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., and Nichols, J. (2016). Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. Stem Cell Reports *6*, 437–446.

Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W., et al. (2017). Epigenetic resetting of human pluripotency. Development *144*, 2748–2763.

Halevy, T., and Urbach, A. (2014). Comparing ESC and iPSC-Based Models for Human Genetic Disorders. Journal of clinical medicine, *3*, 1146–1162.

Hao, J., Li, T.-G., Qi, X., Zhao, D.-F., and Zhao, G.-Q. (2006). WNT/beta-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. Dev. Biol. *290*, 81–91.

Hu, K. (2014). Vectorology and Factor Delivery in Induced Pluripotent Stem Cell Reprogramming.

Stem Cells Dev. *23*, 1301–1315.

Huang, S.-M.A., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature *461*, 614–620.

Ivics, Z.N., Hackett, P.B., Plasterk, R.H., and Izsvák, Z. (1997). Molecular Reconstruction of Sleeping Beauty, a Tc1-like Transposon from Fish, and Its Transposition in Human Cells. Cell *91*, 501–510.

Iwamoto, M., Bjorklund, T., Lundberg, C., Kirik, D., and Wandless, T.J. (2010). A general chemical method to regulate protein stability in the mammalian central nervous system. Chem. Biol. *17*, 981–988.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Jouhilahti, E.-M., Madissoon, E., Vesterlund, L., Töhönen, V., Krjutškov, K., Plaza Reyes, A., Petropoulos, S., Månsson, R., Linnarsson, S., Bürglin, T., et al. (2016). The human PRD-like homeobox gene LEUTX has a central role in embryo genome activation . Development *143*, 3459–3469.

Kim, H., Wu, J., Ye, S., Tai, C.I., Zhou, X., Yan, H., Li, P., Pera, M., and Ying, Q.L. (2013). Modulation of β-catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal. Nat. Commun. *4*, 2403.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895–2902.

Llames, S., Garcia-Perez, E., Meana, A., Larcher, F., and del Rio, M. (2015). Feeder Layer Cell Actions and Applications. Tissue Eng. Part B. Rev. *21*, 345–353.

Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013). CRISPR RNA–guided activation of endogenous human genes. Nat. Methods *10*, 977.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., et al. (2007). Directly Reprogrammed Fibroblasts
Show Global Epigenetic Remodeling and Widespread Tissue Contribution. Cell Stem Cell *1*, 55–

70.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science. *339*, 823–826.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Gottgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell *11*, 491–504.

Morey, L., and Helin, K. (2010). Polycomb group protein-mediated repression of transcription. Trends Biochem. Sci. *35*, 323–332.

Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T., and Saitou, M. (2016). A developmental coordinate of pluripotency among mice, monkeys and humans. Nature *537*, 57–62.

Neri, F., Krepelova, A., Incarnato, D., Maldotti, M., Parlato, C., Galvagni, F., Matarese, F., Stunnenberg, H.G., and Oliviero, S. (2013). Dnmt3L antagonizes DNA methylation at bivalent promoters and favors DNA methylation at gene bodies in ESCs. Cell *155*, 121.

Nichols, J., and Smith, A. (2009). Naive and Primed Pluripotent States. Cell Stem Cell *4*, 487–492.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935–949.

Nusse, R., and Varmus, H. (2012). Three decades of Wnts: a personal perspective on how a scientific field developed. EMBO J. *31*, 2670–2684.

O'Leary, T., Heindryckx, B., Lierman, S., van Bruggen, D., Goeman, J.J., Vandewoestyne, M., Deforce, D., de Sousa Lopes, S.M.C., and De Sutter, P. (2012). Tracking the progression of the human inner cell mass during embryonic stem cell derivation. Nat. Biotechnol. *30*, 278–282.

Ogawa, K., Nishinakamura, R., Iwamatsu, Y., Shimosato, D., and Niwa, H. (2006). Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. Biochem. Biophys. Res. Commun. *343*, 159–166.

Pastor, W.A., Chen, D., Liu, W., Kim, R., Sahakyan, A., Lukianchikov, A., Plath, K., Jacobsen, S.E., and Clark, A.T. (2016). Naive Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. Cell Stem Cell *18*, 323–329.

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat. Methods *10*, 973–976.

Plasterk, R.H., Izsvak, Z., and Ivics, Z. (1999). Resident aliens: the Tc1/mariner superfamily of transposable elements. Trends Genet. *15*, 326–332.

Romito, A., and Cobellis, G. (2016). Pluripotent stem cells: Current understanding and future directions. Stem Cells Int. *2016*.

Roode, M., Blair, K., Snell, P., Elder, K., Marchant, S., Smith, A., and Nichols, J. (2012). Human hypoblast formation is not dependent on FGF signalling. Dev. Biol. *361*, 358–363.

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T.W., Kuoy, E., Langerman, J., Clark, A.T., Jaenisch, R., et al. (2017). Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. Cell Stem Cell *20*, 87–101.

Schmitz, Y., Rateitschak, K., and Wolkenhauer, O. (2013). Analysing the impact of nucleo-cytoplasmic shuttling of beta-catenin and its antagonists APC, Axin and GSK3 on Wnt/beta-catenin signalling. Cell. Signal. *25*, 2210–2221.

Shimoda, M., Kanai-Azuma, M., Hara, K., Miyazaki, S., Kanai, Y., Monden, M., and Miyazaki, J. (2007). Sox17 plays a substantial role in late-stage differentiation of the extraembryonic endoderm in vitro. J. Cell Sci. *120*, 3859–3869.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol. *6*, 2237–2247.

Silva, J., & Smith, A. (2008). Capturing pluripotency. Cell. *132*, 532–536.

Smith, A.G. (2001). Embryo-derived stem cells: of mice and men. Annu. Rev. Cell Dev. Biol. *17*, 435–462.

Sobhani, A., Khanlarkhani, N., Baazm, M., Mohammadzadeh, F., Najafi, A., Mehdinejadiani, S., and Aval, F.S. (2017). Multipotent stem cell and current application. Acta Med. Iran. *55*, 6–23.

Stavridis, M.P., Lunn, J.S., Collins, B.J., and Storey, K.G. (2007). A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. Development *134*, 2889–2894.

Tabar, M.S., Hesaraki, M., Esfandiari, F., Samani, F.S., Vakilian, H., and Baharvand, H. (2015). Evaluating electroporation and lipofectamine approaches for transient and stable transgene expressions in human fibroblasts and embryonic stem cells. Cell J. *17*, 438–450.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. Cell *158*, 1254–1269.

Tanaka, S.S., Kojima, Y., Yamaguchi, Y.L., Nishinakamura, R., and Tam, P.P.L. (2011). Impact of WNT signaling on tissue lineage differentiation in the early mouse embryo. Dev. Growth Differ. *53*, 843–856.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196–199.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell *15*, 471–487.

Theunissen, T.W., Friedli, M., He, Y., Planet, E., O'Neil, R.C., Markoulaki, S., Pontis, J., Wang, H., Iouranova, A., Imbeault, M., et al. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. Cell Stem Cell *19*, 502–515.

Thomas, P., and Smart, T.G. (2005). HEK293 cell line: a vehicle for the expression of recombinant proteins. J. Pharmacol. Toxicol. Methods *51*, 187–200.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. Science. *282*, 1145.

Töhönen, V., Katayama, S., Vesterlund, L., Jouhilahti, E.-M., Sheikhi, M., Madissoon, E., Filippini-Cattaneo, G., Jaconi, M., Johnsson, A., Burglin, T.R., et al. (2015). Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. Nat. Commun. *6*, 8207.

Vad-Nielsen, J., Lin, L., Bolund, L., Nielsen, A.L., and Luo, Y. (2016). Golden Gate Assembly of

CRISPR gRNA expression array for simultaneously targeting multiple genes. Cell. Mol. Life Sci. *73*, 4315–4325.

Valenta, T., Hausmann, G., and Basler, K. (2012). The many faces and functions of beta-catenin. EMBO J. *31*, 2714–2736.

Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J. Cell Sci. *118*, 4495–4509.

Wang, W., Lin, C., Lu, D., Ning, Z., Cox, T., Melvin, D., Wang, X., Bradley, A., and Liu, P. (2008). Chromosomal transposition of PiggyBac in mouse embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *105*, 9290–9295.

Ware, C.B., Wang, L., Mecham, B.H., Shen, L., Nelson, A.M., Bar, M., Lamba, D.A., Dauphin, D.S., Buckingham, B., Askari, B., et al. (2009). Histone Deacetylase Inhibition Elicits an Evolutionarily Conserved Self-Renewal Program in Embryonic Stem Cells. Cell Stem Cell *4*, 359–369.

Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, S., Nishikawa, S.I., Muguruma, K., et al. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat. Biotechnol. *25*, 681–686.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. Nat. Rev. Mol. Cell Biol. *17*, 155–169.

Weltner, J. (2018). Novel Approaches for Pluripotent Reprogramming. (Doctoral dissertation). Retrieved from //helda.helsinki.fi/bitstream/handle/10138/237188/Nov elApp.pdf

Weltner, J., Balboa, D., Katayama, S., Bespalov, M., Krjutškov, K., Jouhilahti, E.-M., Trokovic, R., Kere, J., and Otonkoski, T. (2018). Human pluripotent reprogramming with CRISPR activators. Nat. Commun. *9*, 2643.

Wobus, A.M., and Boheler, K.R. (2005). Embryonic Stem Cells: Prospects for Developmental Biology and Cell Therapy. Physiol. Rev. *85*, 635–678.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat. Cell Biol. *13*, 838–845.

Xu, Z., Robitaille, A.M., Berndt, J.D., Davidson, K.C., Fischer, K.A., Mathieu, J., Potter, J.C.,

Ruohola-Baker, H., and Moon, R.T. (2016). Wnt/β-catenin signaling promotes self-renewal and inhibits the primed state transition in naïve human embryonic stem cells. Proc. Natl. Acad. Sci. *113*, E6382–E6390.

Yang, P., Humphrey, S.J., Cinghu, S., James, D.E., Mann, M., Yang, P., Humphrey, S.J., Cinghu, S., Pathania, R., and Oldfield, A.J. (2019). Multi-omic Profiling Reveals Dynamics of the Phased Progression of Pluripotency Article Multi-omic Profiling Reveals Dynamics of the Phased Progression of Pluripotency. Cell Syst. *8*, 427–445.

Yang, Y., Liu, B., Xu, J., Wang, J., Wu, J., Shi, C., Xu, Y., Dong, J., Wang, C., Lai, W., et al. (2017). Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. Cell *169*, 243–257.

Ye, S., Liu, D., and Ying, Q.L. (2014). Signaling pathways in induced naïve pluripotency. Curr. Opin. Genet. Dev. *28*, 10–15.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281–292.

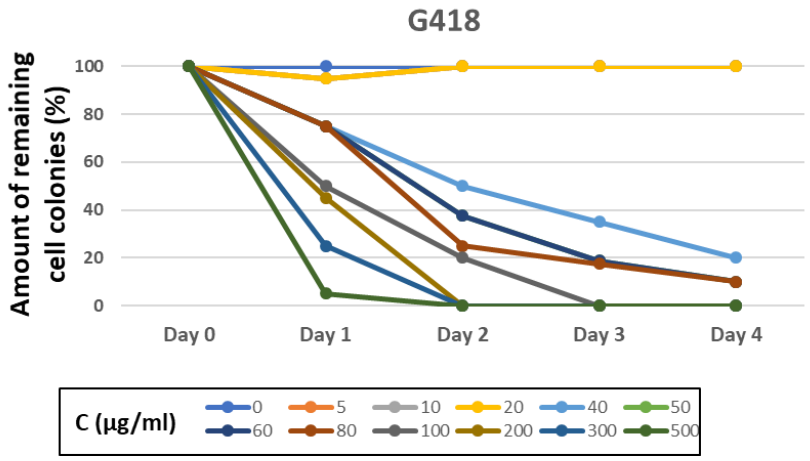
Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

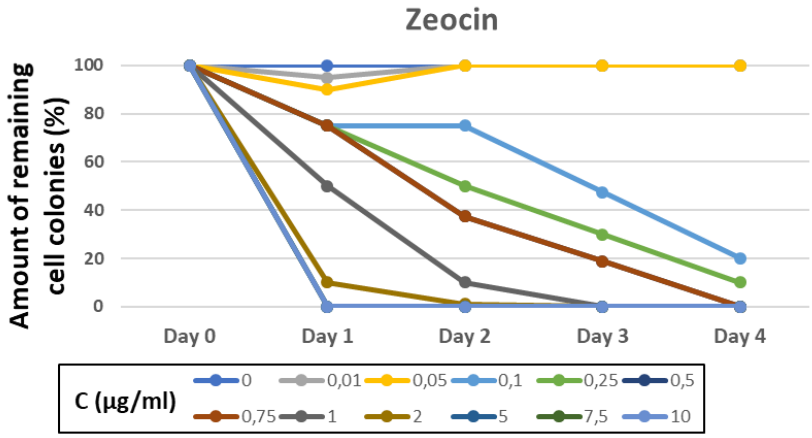
Zhong, Y., and Holland, P.W.H. (2011). The dynamics of vertebrate homeobox gene evolution: gain and loss of genes in mouse and human lineages. BMC Evol. Biol. *11*, 169.

Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., Blau, C.A., Horwitz, M.S., Hockenbery, D., Ware, C., et al. (2012). HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. EMBO J. *31*, 2103–2116.

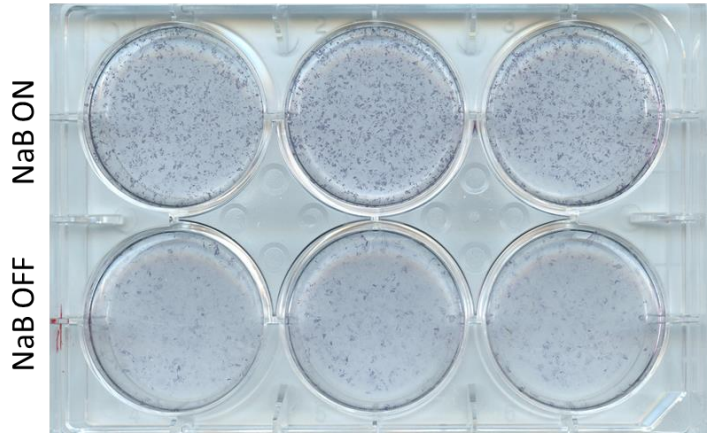
Zimmerlin, L., Park, T.S., Huo, J.S., Verma, K., Pather, S.R., Talbot, C.C., Agarwal, J., Steppan, D., Zhang, Y.W., Considine, M., et al. (2016). Tankyrase inhibition promotes a stable human naïve pluripotent state with improved functionality. Dev. *143*, 4368–4380.

8. Appendices

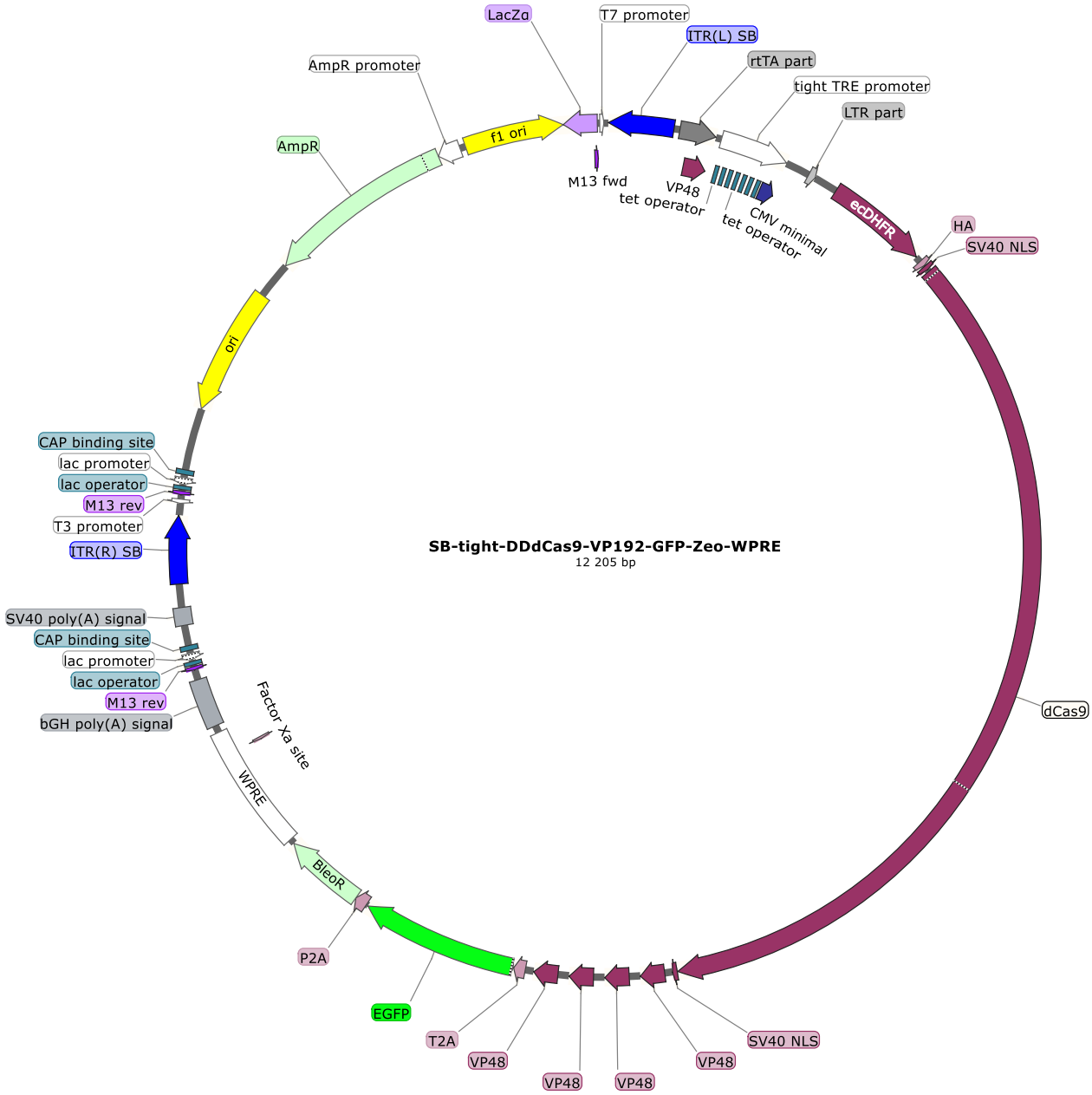




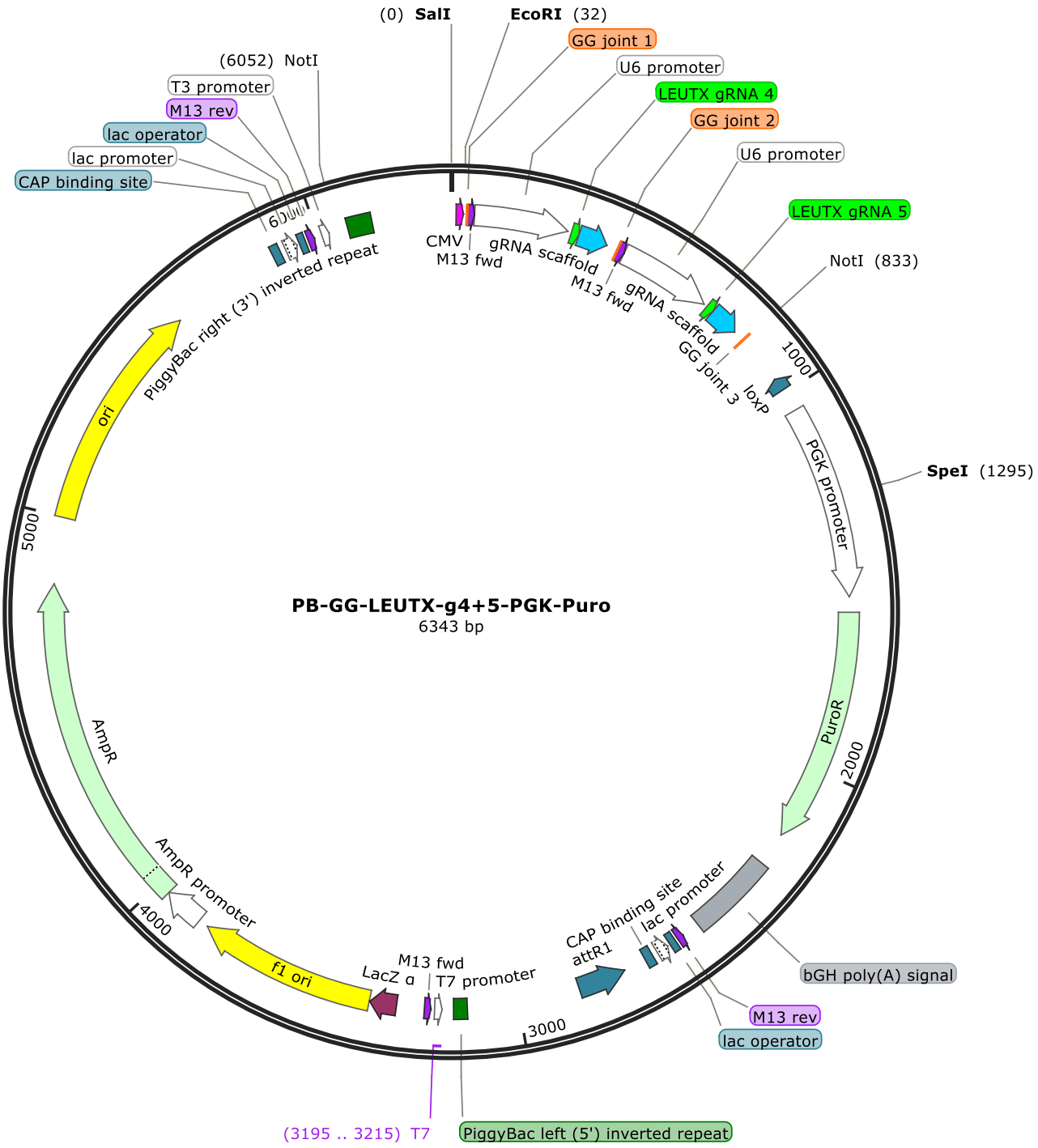
Supplementary Figure 1. Four day kill-curve assays for optimizing the neomycin and zeocin antibiotic selection concentrations on H9 PSCs



Supplementary Figure 2. AP-stained iPSC colonies formed by HFFs used for NaB viability testing. The upper three wells were induced with 0.25 mM NaB



Supplementary Figure 3. The plasmid map of SB-tight-DDdCas9-VP192-GFP-Zeo-WPRE containing the DDdCas9VP192 activator sequence



Supplementary Figure 4. The plasmid map of PB-GG-LEUTX-g4+5-PGK-Puro containing the gRNAs for LEUTX used in reversions